Cybersécurité

Chiffrement et hachage des données

Prof. M. LAROUSSI

Définitions et explication

- En cybersécurité, le chiffrement désigne **la conversion des données** depuis un format lisible dans un format codé.
- Les données chiffrées ne peuvent être lues ou traitées qu'après leur déchiffrement.
- Le chiffrement **est l'élément fondamental de la sécurité** des données.
- C'est le moyen **le plus simple et le plus efficace de s'assurer que les informations du système informatique ne peuvent être ni volées ni lues** par quelqu'un qui souhaite les utiliser à des fins malveillantes.
- Le chiffrement à des fins de sécurité des données **est largement utilisé par les particuliers et les grandes entreprises** pour protéger les informations des utilisateurs envoyées entre un navigateur et un serveur.
- Ces informations peuvent inclure tout type de renseignements, des données de paiement aux informations personnelles.
- Le logiciel de chiffrement des données, **également appelé algorithme de chiffrement ou code**, est utilisé pour développer un schéma de chiffrement qui, en théorie, ne peut être rompu que par une **exceptionnelle puissance de calcul**.

Définitions et explication

- Le chiffrement désigne la conversion de texte brut lisible par les hommes en texte incompréhensible, appelé texte chiffré.
- Cette conversion correspond à la capture de données lisibles et à leur modification en données apparemment aléatoires.
- Le chiffrement implique l'utilisation d'une clé cryptographique, c'est-à-dire un ensemble de valeurs mathématiques convenues par l'expéditeur et le destinataire.
- Le destinataire utilise la clé pour déchiffrer les données et les transformer en texte brut lisible.
- Plus la clé cryptographique est complexe, plus le chiffrement est sécurisé car les tiers sont moins susceptibles de déchiffrer des données via des attaques par force brute (par exemple en tentant des numéros aléatoires jusqu'à deviner la combinaison correcte).
- Le chiffrement est également utilisé pour protéger les mots de passe. Les méthodes de chiffrement des mots de passe brouillent votre mot de passe qui devient illisible pour les cybercriminels.

Méthodes de chiffrement

Les deux principales techniques de chiffrement sont le chiffrement **symétrique** et **asymétrique**.

Ces noms font référence à la clé, qui peut être la même ou non pour le chiffrement et le déchiffrement :

Clés de chiffrement symétrique

- Appelé chiffrement à clé privée.
- La clé utilisée pour encoder est la même que celle utilisée pour décoder, ce qui convient parfaitement pour les utilisateurs individuels et les systèmes fermés.
- La clé doit être envoyée au destinataire, ce qui augmente le risque de compromission si elle est interceptée par un tiers (un cybercriminel, par exemple).
- Cette méthode est plus rapide que la méthode asymétrique.

Clés de chiffrement asymétrique

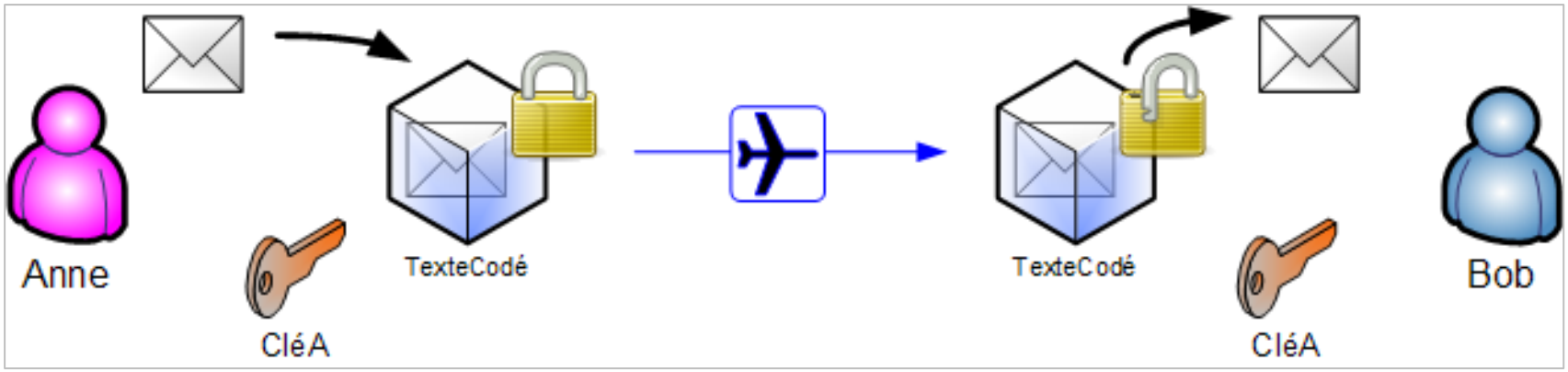
- cette méthode utilise deux clés différentes (publique et privée) mathématiquement reliées.
- Concrètement, les clés se composent uniquement de grands nombres qui ont été couplés entre eux mais ne sont pas identiques, d'où le terme asymétrique.
- La clé privée est tenue secrète par le propriétaire et la clé publique est soit partagée parmi les destinataires autorisés, soit mise à disposition du public à grande échelle.

Chiffrement symétrique

- Les algorithmes de chiffrement symétrique **se fondent sur une même clé pour chiffrer et déchiffrer un message**.

- L'un des problèmes de cette technique est que la clé, qui doit rester totalement confidentielle, doit être transmise au correspondant de façon sûre.

- La mise en œuvre peut s'avérer difficile, surtout avec un grand nombre de correspondants car il faut autant de clés que de correspondants.



Algorithme de chiffrement symétrique

- Quelques algorithmes de chiffrement symétrique très utilisés :

- Chiffre de Vernam (le seul offrant une sécurité théorique absolue, à condition que la clé ait au moins la même longueur que le message, qu'elle ne soit utilisée qu'une seule fois à chiffrer et qu'elle soit totalement aléatoire).

  - DES
  - 3DES
  - AES
  - RC4
  - RC5
  - MISTY1

Algorithme de chiffrement symétrique

On distingue deux catégories de chiffrement symétrique :

- Le **chiffrement par bloc** (**Block Cipher**) est une des deux grandes catégories de chiffrements modernes en cryptographie symétrique, l'autre étant le chiffrement par flot.

- C'est la méthode la plus utilisée aujourd'hui, avec l'algorithme AES (Advanced Encryption Standard), c'est le successeur de DES (D: Data)

- L'algorithme le plus largement utilisé

- La principale différence vient du découpage des données en blocs de taille généralement fixe.

- La taille de bloc est comprise entre 32 et 512 bits, dans le milieu des années 1990 le standard était de 64 bits mais depuis 2000 et le concours **AES le standard est de 128 bits**

Algorithme de chiffrement symétrique

Plusieurs modes existent, certains sont plus vulnérables que d'autres :

- Dictionnaire de codes (Electronic Code Book, ECB)
- Enchaînement des blocs (Cipher Block Chaining, CBC)
- Chiffrement à rétroaction (Cipher Feedback, CFB)
- Chiffrement à rétroaction de sortie (Output Feedback, OFB)
- Chiffrement basé sur un compteur (CounTeR, CTR)
- Chiffrement avec vol de texte (CipherText Stealing, CTS)
- …

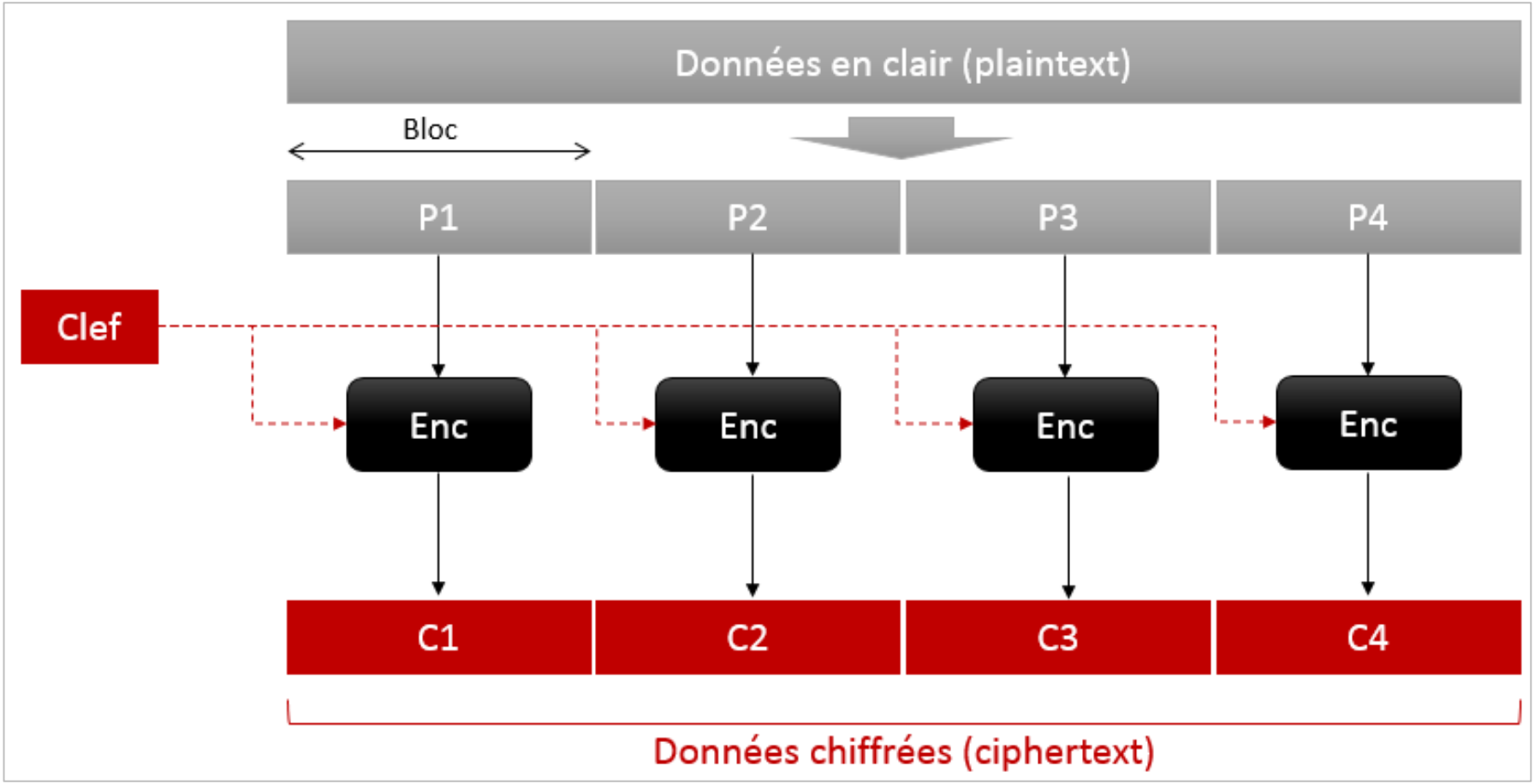
Algorithme de chiffrement symétrique

Le chiffrement ECB (Electronic Code Book)

- Dans le mode de chiffrement par bloc le plus simple, appelé **ECB**, on réalise le chiffrement de chaque bloc séparément et on réassemble tous les blocs en sortie pour former le texte chiffré.
- C'est la méthode de chiffrement la plus simple à utiliser
- Le mode ECB comporte un défaut de taille : **tous les blocs en clair identiques vont donner le même bloc chiffré.**
- Le mode ECB ne respecte pas l'intégrité des données.
- Un attaquant peut remplacer certains blocs chiffrés par d'autres blocs chiffrés du message, ou permuter deux blocs, sans que le destinataire s'en aperçoive.
- Imaginons que le message chiffré soit le montant d'une transaction électronique, et que l'attaquant arrive à permuter deux chiffres.

Algorithme de chiffrement symétrique

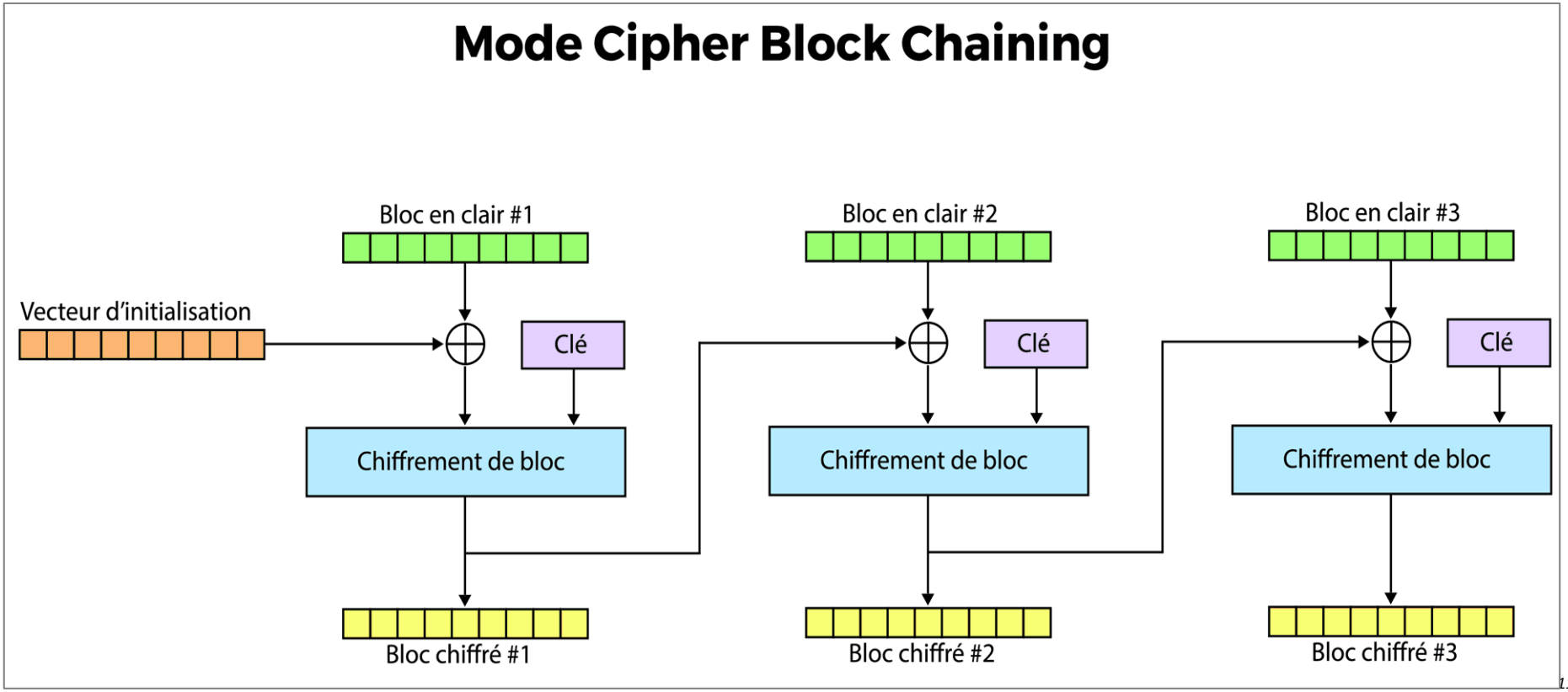
Le chiffrement ECB (Electronic Code Book)



Le chiffrement CBC (Cipher Block Chaining)

- Ce mode consiste à effectuer l'opération XOR entre le bloc clair actuel et le bloc chiffré précédent, avant de chiffrer le bloc actuel.

- Pour le premier bloc, il n'y a pas de bloc précédent. On effectue l'opération XOR avec une variable aléatoire appelée **vecteur d'initialisation** (IV), et on transmet l'IV en clair avec le texte chiffré.



Chiffrement par flot (à flux)

- Le chiffrement par flot (Stream Cipher) est une des deux grandes catégories de chiffrements modernes (Chiffrements par flux et chiffrement pas bloc), **utilisant une seule clé**.

- Un chiffrement par flot arrive à traiter les données de longueur quelconque **et n'a pas besoin de les découper**.

- Un chiffrement par flot se présente souvent sous la forme d'un générateur de nombres pseudo-aléatoires avec lequel on opère un XOR entre un bit à la sortie du générateur et un bit provenant des données.

- Il est très utilisé dans le contexte de chiffrement des communications entre objet connectés.

- Des algorithmes de chiffrement par flot :

  - *A5/1, algorithme publié en 1994, utilisé dans les téléphones mobiles de type GSM pour chiffrer la communication par radio entre le mobile et l'antenne-relais la plus proche,*
  - *RC4, le plus répandu, conçu en 1987 par Ronald Rivest, utilisé notamment par le protocole WEP du WiFi*
  - *E0 utilisé par le protocole Bluetooth*

Chiffrement par flot (à flux)

Le principe du chiffrement par flot est de chiffrer une suite de caractères (ou octets ou mots-machine) un à la fois, à l'aide d'une transformation qui varie au fur et à mesure du texte.

Au contraire, le chiffrement par bloc utilise une transformation fixe, sur des blocs plus gros, typiquement 64 ou 128 bits.

La formule à appliquer :

Soit $\oplus$ l'opération booléenne XOR, et $G_m : \{0,1\}^n \to \{0,1\}^m$ un générateur pseudo-aléatoire. Ce mécanisme construit sur le principe du masque jetable utilise la propriété involutive du OU exclusif.

- Chiffrement du message $M \in \{0,1\}^m$ avec la clé $K \in \{0,1\}^n$: $M \oplus G_m(K) = C$.
- Déchiffrement du message C avec la clé K : $C \oplus G_m(K) = (M \oplus G_m(K)) \oplus G_m(K) = M \oplus (G_m(K) \oplus G_m(K)) = M$

Hachage : Définitions et principes

- Le hachage est la transformation d'une chaîne de caractères en valeur, clé ou signature de longueur fixe, *généralement plus courte*, représentant la chaîne d'origine.
- Le hachage est notamment employé pour indexer et récupérer les éléments d'une base de données, comme les mots de passe.
- Il est en effet plus rapide de trouver l'élément d'après la clé de hachage réduite plutôt qu'à l'aide de la valeur d'origine.
- Cette fonction est également utilisée dans de nombreux algorithmes de chiffrement.

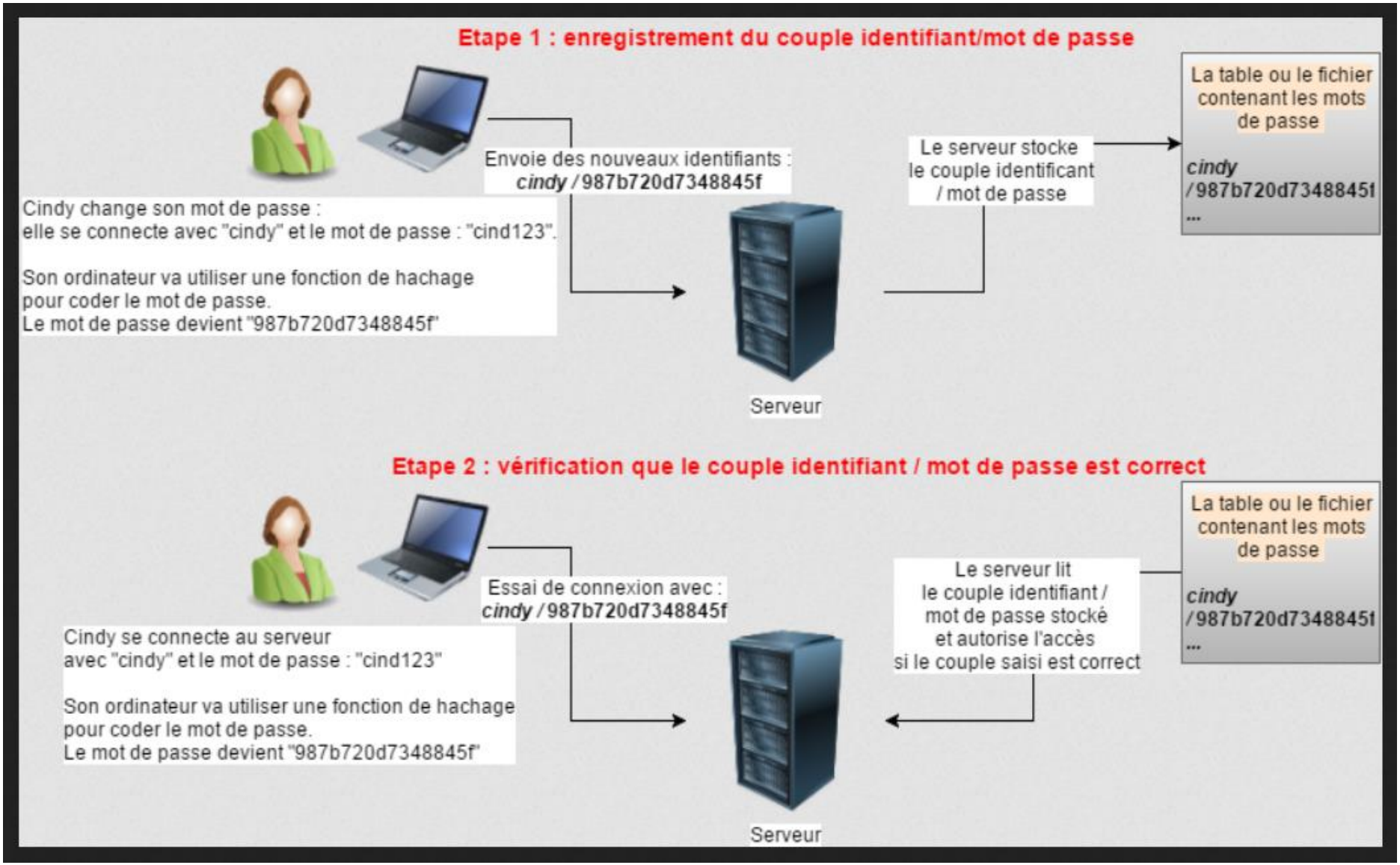
Principe du hachage

Dans le hachage, le mot de passe (Data) n'est pas envoyé au serveur, mais une signature du mot de passe.

Le serveur ne va pas enregistrer le mot de passe mais enregistrera cette signature.

Lorsque l'utilisateur se connectera, le serveur ne va pas vérifier si le mot de passe est identique, mais il va vérifier que la signature du mot de passe saisi est bien la même que la signature du mot de passe enregistré.

Hachage : Définitions et principes



Fonctions de hachage

- La longueur de la signature (Clé) **doit être toujours la même** (quelque soit la longueur des données en entrée.)
- Il n'est pas possible de trouver les données originales à partir des empreintes : Les fonctions de hachage ne fonctionnent que dans un seul sens : Le processus inverse est interdit.
- Il ne doit pas être possible de prédire une signature. (Il n'est pas possible d'essayer d'imaginer ce que pourrait être la signature en examinant les données)
- Et enfin, évidemment pour des données différentes : les signatures doivent être différentes.

Exemples de fonctions de hachage connues

MD5

Vous en avez forcément entendu parlé. Cette fonction de hachage est toujours très utilisée bien qu'au niveau de la sécurité, il est recommandé de passer à des versions plus robustes car des suites de collisions ont été trouvées.

Cette fonction renvoie une empreinte de 128 bits.

SHA1

Était la fonction remplaçante de MD5 car elle produisait des empreintes 160 bits et avec impossibilité de trouver des collisions … jusqu'en 2004-2005, date à laquelle des attaques ont prouvé des possibilités de générer des collisions. Depuis, cette date il n'est plus conseillé d'utiliser la fonction SHA1.

Les certificats numériques utilisant SHA1 ne sont plus plus valides au 31 décembre 2016.

SHA2

SHA256 et SHA512 sont 2 des grands standards utilisés actuellement, car il n'y a pas à ce jour d'attaques ayant trouvé des failles de sécurité sur ces fonctions de hachage.

Les signatures générées sont respectivement de 256 et 512 bits.

Certificat numérique, c'est quoi ?

Qu'est-ce qu'un certificat numérique ?

- Un certificat numérique est un type de fichier utilisé pour associer des paires de clés cryptographiques à des entités telles que des sites Web, des individus ou des organisations.
- Si la confiance du public est requise, un Autorité de certification (CA) comme SSL.com valide les identifie et les associe à des paires cryptographiques via des certificats numériques.
- La paire de clés mentionnée se compose d'une clé publique et d'une clé privée. La clé publique est incluse dans le certificat, tandis que la clé privée est sécurisée.
- Le propriétaire de la clé privée peut ensuite l'utiliser pour signer des documents, et la clé publique peut être utilisée pour vérifier la validité de ces signatures.
- Des tiers peuvent également utiliser la clé publique pour envoyer des informations chiffrées, que seul le propriétaire de la clé privée peut chiffrer.

Signatures numériques

- La **signature électronique** est un outil numérique et juridique qui permet de faire signer tous les documents électroniques en ligne de manière **rapide** et **sécurisée**.
- La signature électronique est un mécanisme permettant de garantir l'intégrité d'un document électronique et d'en authentifier l'auteur, par analogie avec la signature manuscrite d'un document papier.
- Elle a **la même valeur légale qu'une signature manuscrite**.
- Elle se différencie de la signature écrite par le fait qu'elle n'est pas visuelle, mais correspond à une suite de caractères.

Selon l'article 1366 du Code civil :

« L'écrit électronique a la même force probante que l'écrit sur support papier, sous réserve que puisse être dûment identifiée la personne dont il émane et qu'il soit établi et conservé dans des conditions de nature à en garantir l'intégrité. »

Signature numérique

- La signature électronique s'appuie sur un système d'horodatage, une clé chiffrée d'authentification et un certificat électronique.

- Ces procédés permettent de générer une empreinte numérique propre à chaque document signé de manière électronique, garantissant leur authenticité.

Un signature électronique doit être :

Authentique : L'identité du signataire doit pouvoir être retrouvée de manière certaine

Infalsifiable : Une personne ne peut pas se faire passer pour un autre

Non réutilisable : La signature fait partie du document signé et ne peut être déplacée sur un autre document

Inaltérable : Une fois que le document est signé, on ne peut plus le modifier

Irrévocable : La personne qui a signé ne peut le contester

Chiffrement asymétrique

- **Définitions**
- **Fonctionnement**
- **Algorithmes**
- **Avantages et inconvénients**
- **Utilisation**