12 : Introduction à la Cybersécurité

I - Introduction

La **cybersécurité** est la **pratique** consistant à **protéger** les systèmes, les réseaux et les programmes **contre les attaques numériques**.

II - La nécessité de la Cybersécurité

La **Cybersécurité** s'appuie sur le modèle **CID** (**Confidentialité**, **Intégrité** et **Disponibilité**) :

Confidentialité : Protéger l'accès aux données sensibles contre les personnes non autorisées. Elle garantit l'anonymat des données en limitant l'accès par le **chiffrement** (**cryptographie**) ;

Intégrité : Assurer que les données ne sont ni altérées ni corrompues. Elle garantit que l'information est exacte et fiable ;

Disponibilité : Garantir que les personnes autorisées peuvent accéder à l'information.

Remarque : Au-delà des 3 critères **DIC** (**disponibilité**, **intégrité** et **confidentialité**), un quatrième critère est parfois ajouté : la **traçabilité** (aussi appelés critères **DICT**). Si des données ont été modifiées ou supprimées, il peut être important de pouvoir identifier d'où provient cette modification.

III - La confidentialité des données

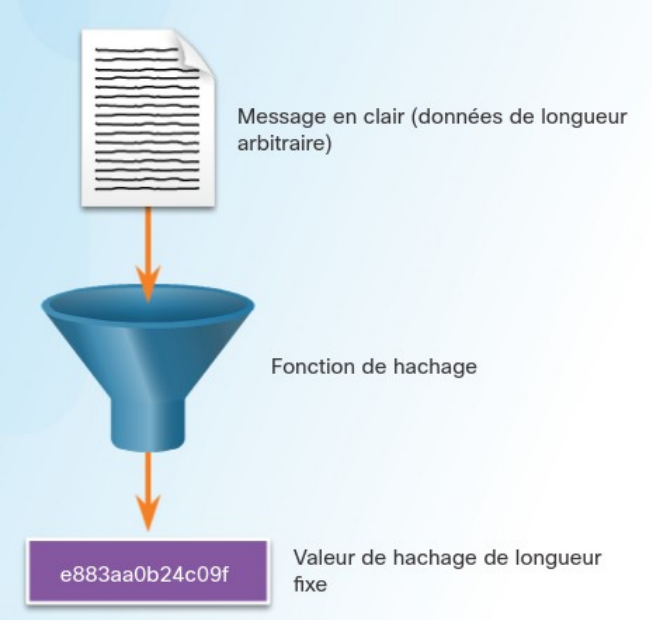
La **confidentialité des données** est réalisée à l'aide :

- de l'**hygiène numérique** qui consiste à avoir un **comportement** permettant de **maintenir les données sensibles en sécurité** et de les **protéger contre les cyberattaques et le vol** (avoir des **mots de passe robustes**, faire des **mises à jour régulières** et être **vigilant en ligne**) ;
- de la **cryptographie** qui consiste à **chiffrer des données** pour les rendre **confidentielles**. Voir **Fiche 13**.

IV - L'intégrité des données

Une **somme de contrôle** est utilisée pour vérifier l'intégrité des fichiers ou des chaînes de caractères après leur transfert d'un périphérique à un autre dans votre réseau local ou sur Internet.
Les **sommes de contrôle** sont calculées grâce à des fonctions de **hachage**. Parmi les sommes de contrôle les plus courantes, il y a **MD5**, **SHA-1**, **SHA-256** et **SHA-512**.

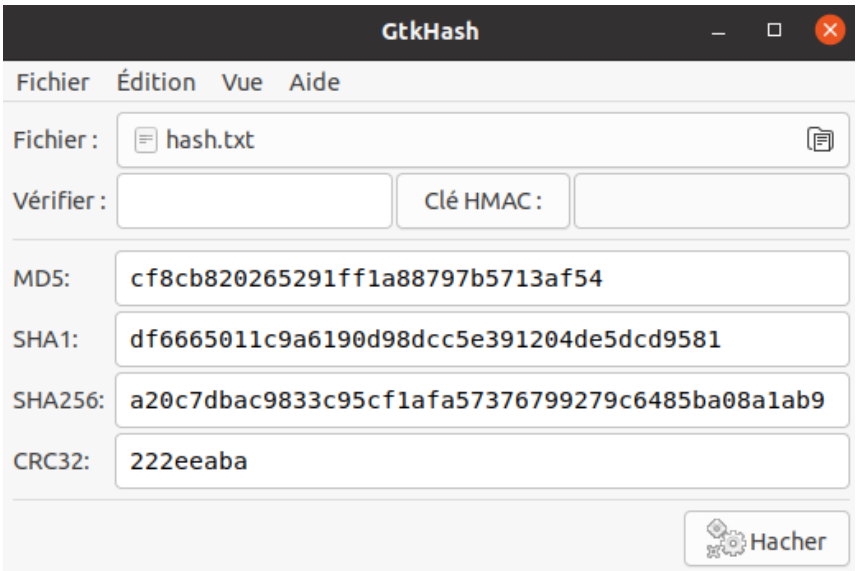
Comme le montre la figure ci-dessous, une fonction de **hash** utilise un algorithme mathématique pour transformer les données en une **valeur de longueur fixe** qui représente les données. La valeur hachée est simplement présente pour comparaison.



Il est impossible d'extraire directement les données d'origine à partir de la valeur hachée. Par exemple, si vous avez oublié votre mot de passe, vous ne pourrez pas le récupérer à partir de la valeur hachée. Il vous faut réinitialiser le mot de passe.

Remarque : Afin d'améliorer l'efficacité du hachage, on peut ajouter un **sel** (**salt**). c'est une donnée informatique qui n'est pas secrète, souvent aléatoire, qui peut être ajoutée au message avant hachage.

On peut générer des **sommes de contrôle** avec le logiciel **HashCalc** sous Windows ou **GtkHash** sous **Ubuntu** :



V - La disponibilité des données

La **disponibilité** garantit que les **utilisateurs autorisés** d'un système ont un **accès rapide** et **ininterrompu** aux **informations** contenues dans ce système, ainsi qu'au réseau.

Cela implique de **mettre en place des mesures** pour **minimiser les temps d'arrêt** et assurer la **continuité des opérations**, même en cas de défaillances matérielles, d'attaques ou de catastrophes naturelles. Des stratégies telles que les **sauvegardes régulières**, les **systèmes redondants** et les **plans de reprise après sinistre** sont essentielles pour maintenir une **haute disponibilité**.

Voici les **méthodes** qui permettent cette **haute disponibilité** :

- **Sauvegardes** : La sauvegarde permet de protéger les données contre divers risques tels que les pannes matérielles, les erreurs humaines, les attaques de logiciels malveillants, les catastrophes naturelles, etc.. La sauvegarde n'est utile que si les données peuvent être restaurées de manière fiable et rapide en cas de besoin.
- **Répartition équitable** : Communément appelée **équilibrage des charges**, la répartition équitable permet de distribuer la charge (demandes de fichiers, acheminement des données, etc.) de manière à ce qu'aucun appareil ne soit trop sollicité.
- **Redondance** : La redondance fait référence aux systèmes qui sont soit dupliqués, soit basculés vers d'autres systèmes en cas de dysfonctionnement. On appelle « **basculement** » le processus de reconstruction d'un système ou de passage à d'autres systèmes lorsqu'une défaillance est détectée. Dans le cas d'un **serveur**, lorsqu'une défaillance est détectée, le serveur bascule vers un serveur redondant. Si, l'environnement exige un niveau de disponibilité élevé, les serveurs doivent être regroupés en clusters.

La **redondance** et le **basculement** s'appuient sur la **tolérance aux pannes**. La tolérance aux pannes est la capacité d'un système à rester opérationnel en cas de défaillance d'un composant critique, par exemple un lecteur de disque, tombe en panne. La technologie **RAID (Redundant Array of Independent Disks)** utilise plusieurs disques pour assurer la tolérance aux pannes :

- **RAID 0** (entrelacement de disques – striping) : Répartition des données sur plusieurs disques sans tolérance aux pannes,
- **RAID 1** (disques en miroir) : Les données sont dupliquées,
- **RAID 3 ou 4** (entrelacement de disques avec parité dédiée) : Un disque supplémentaire est nécessaire pour stocker les contrôles de parité,
- **RAID 5** (entrelacement de disques avec parité distribuée) : Reconstruction des fichiers en cas de panne d'un des disques. Nécessite l'emploi d'au moins trois disques,
- **RAID 6** (entrelacement de disques avec double parité) : Evolution du **RAID 5**.