Algorithmes de hashage

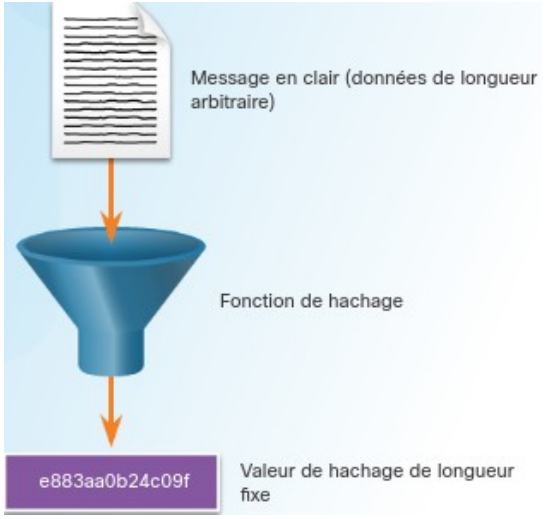
Qu'est-ce que le hash ?

Les utilisateurs du système informatique doivent avoir la garantie que leurs données ne subiront aucune modification, qu'elles soient au repos ou en transit. Le hash est un outil qui assure l'intégrité des données en prenant des données binaires (le message) et en générant une représentation de longueur fixe, appelée valeur de hash ou condensé de message, comme illustré sur cette figure.
L'outil de hash utilise une fonction de hash cryptographique pour vérifier et garantir l'intégrité des données. Il peut également vérifier l'authentification. Les fonctions de hash remplacent les clés de cryptage ou le mot de passe en clair, car il s'agit de fonctions unidirectionnelles.
Cela signifie que si un mot de passe est hashé avec un algorithme de hash spécifique, le condensé de hash obtenu sera toujours le même.

Le qualificatif « unidirectionnel » est utilisé dans la mesure où, avec les fonctions de hash, il est impossible, sur le plan de traitement, que deux ensembles de données différents génèrent une sortie ou un condensé de hash identique.

Chaque fois que les données sont modifiées ou altérées, la valeur de hash change également.

C'est la raison pour laquelle les valeurs de hash cryptographiques sont souvent désignées sous le nom d'empreintes numériques.

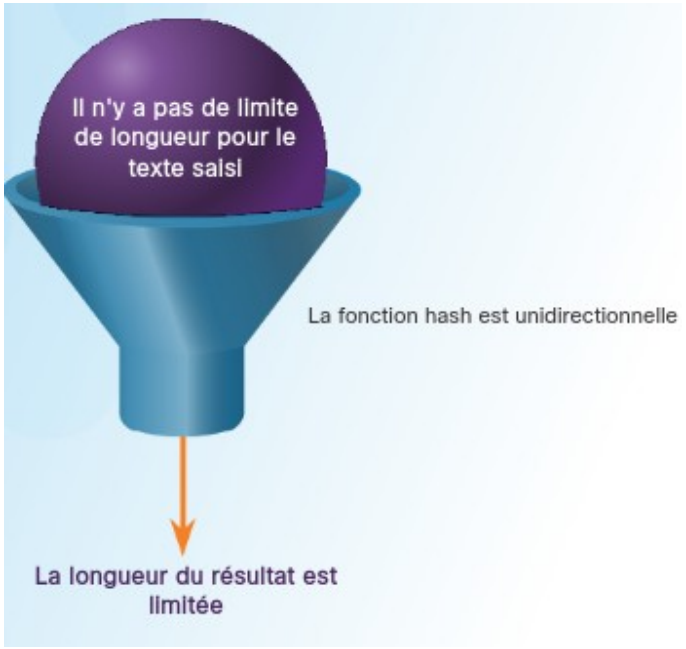


Elles peuvent détecter les fichiers de données en double, les changements de version du fichier et les applications similaires. Ces valeurs constituent une protection contre toute modification accidentelle ou intentionnelle des données, et contre leur corruption accidentelle.
Le hash s'avère également très efficace.
Un fichier volumineux ou le contenu d'un disque dur entier donne comme résultat une valeur de hash de même taille.

Algorithmes de hashage

Propriétés de hash

Le hash est une fonction mathématique unidirectionnelle relativement simple à calculer, mais extrêmement difficile à inverser. La mouture du café est une opération qui illustre parfaitement une fonction unidirectionnelle. Il est facile de moudre des grains de café, mais il est quasiment impossible de les reconstituer ensuite.
Une fonction de hash cryptographique possède les propriétés suivantes :
- Il n'y a pas de limite de longueur pour le texte saisi.
- La longueur du résultat est fixe.
- La fonction de hash est unidirectionnelle et irréversible.



- Deux valeurs d'entrée différentes donneront pratiquement toujours deux valeurs de hash différentes .

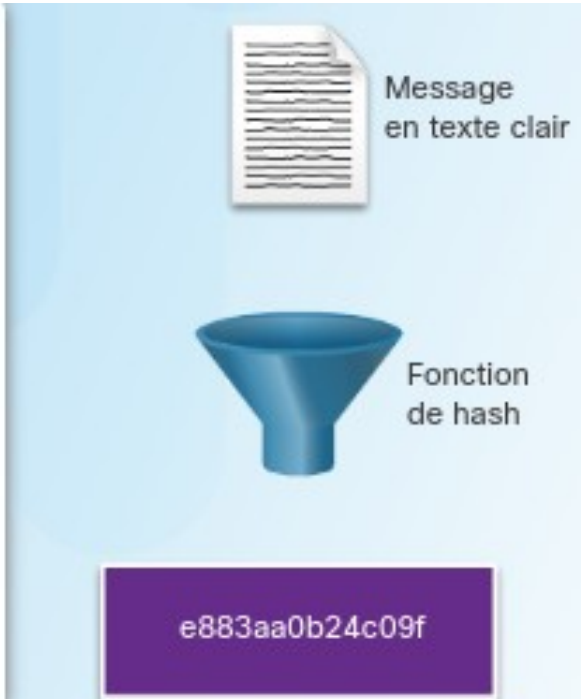
Algorithmes de hashage

Algorithmes de hash modernes

De nombreux algorithmes de hash modernes sont largement utilisés de nos jours. Les plus populaires sont MD5 et SHA.
Algorithme MD5 (Message Digest 5)
C'est à Ron Rivest que l'on doit le développement de MD5, un algorithme de hachage utilisé aujourd'hui par plusieurs applications Internet. MD5 est une fonction unidirectionnelle qui permet de calculer facilement un hash à partir des données d'entrée spécifiées. En revanche, calculer les données d'entrée en connaissant uniquement une valeur de hash s'avère très difficile.
L'algorithme MD5 génère une valeur de hash de 128 bits. Le malware Flame a compromis la sécurité de MD5 en 2012. Les créateurs du malware Flame ont utilisé une collision MD5 pour falsifier un certificat de signature de code Windows. Cliquez ici pour lire un article consacré à l'attaque par collision du malware Flame.
Secure Hash Algorithm (SHA)
L'Institut national des normes et de la technologie (NIST) des États-Unis a développé SHA, l'algorithme spécifié dans la norme SHS (Secure Hash Standard). La publication de l'algorithme SHA-1 date de 1994. SHA-2 a remplacé SHA-1 en ajoutant quatre fonctions de hash qui composent la famille SHA :
SHA-224 (224 bits)
SHA-256 (256 bits)
SHA-384 (384 bits)
SHA-512 (512 bits)

MD5 (Message Digest 5)

MD5 est une séquence complexe d'opérations binaires simples effectuées sur les données d'entrée afin de produire un condensé de message de type hash de 128 bits.



FIN