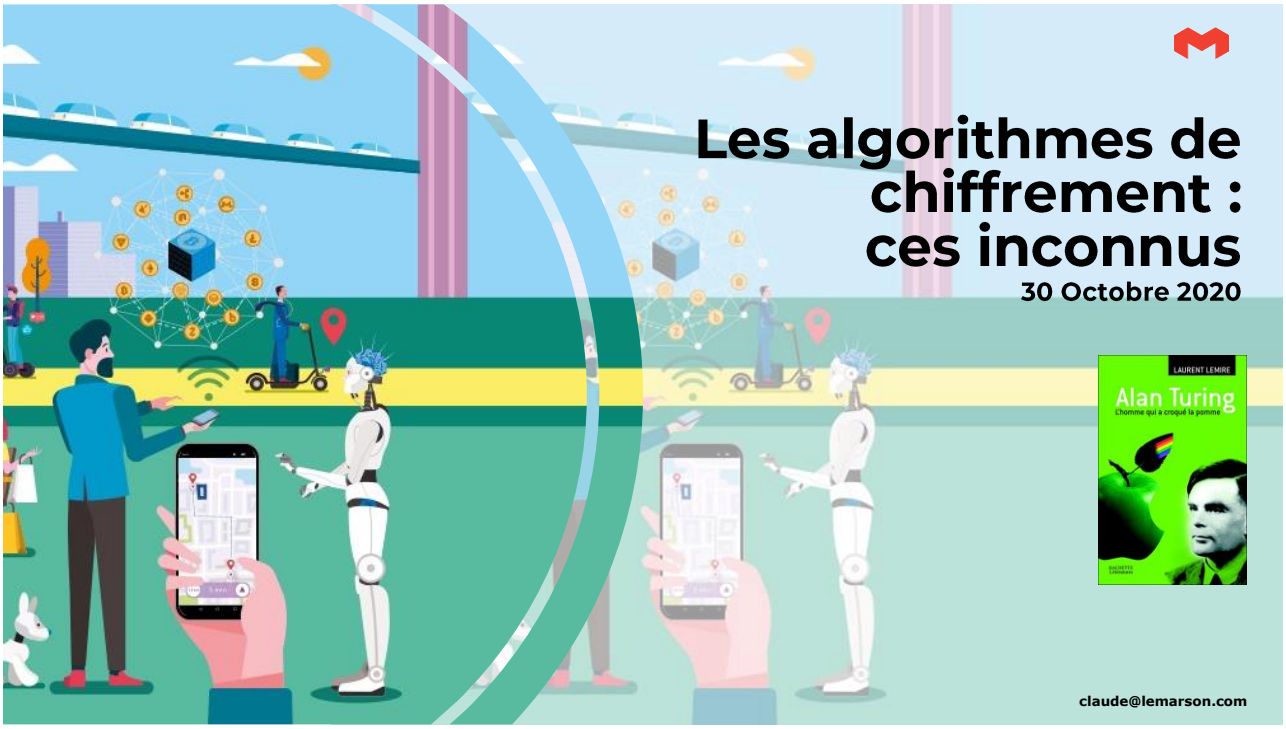


# Les algorithmes de chiffrement : ces inconnus

**30 Octobre 2020**



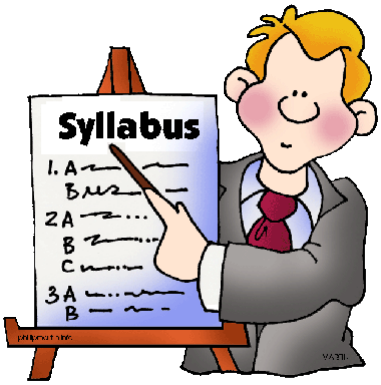

claude@lemarson.com

Un marché de 9,8 G$ en 2020 à 20,1 G$

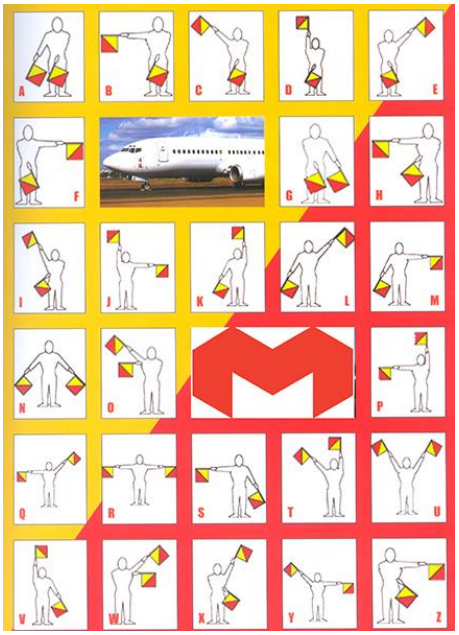
## Sommaire

**Les algorithmes de chiffrement : ces inconnus**





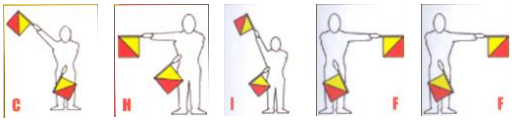

Les technologies cryptographiques seront au cœur des affrontements sécuritaires inter-entreprises et inter-états

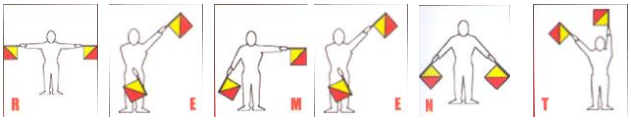


# Ne pas confondre chiffrement avec codage

Le codage n'a pas pour objet de cacher quoi que ce soit. Il peut être associé à une opération de chiffrement.

## CHIFFREMENT





La première ligne de **sémaphores Chappe** (télégraphe) fut établie entre Paris et Lille en 1792. Elle fut utilisée pour transmettre les dépêches de la guerre entre la France et l'Autriche. Le premier symbole d'un message pour Lille traversa 193 km à travers 15 stations en seulement 9 minutes.



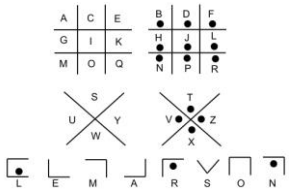
# Un peu d'histoire

## (2 400 ans quand même...)

- L'histoire de la cryptographie remonte aux grecs de Sparte, 404 av JC, avec un système de bande enroulée autour d'un bâton
- César l'utilisait pour communiquer avec ses généraux, en effectuant des permutations de lettres
  - Difficulté : s'il n'y a que 26 lettres, il y a 26! permutations, soit environ 4 milliards de milliards de combinaisons : impossible de trouver une méthode mathématique pour détecter les bonnes combinaisons
- Il faut attendre le savant arabe Al Kindi, au IXème siècle qui se fonde sur la fréquence d'apparition des lettres, différente selon les langues et propose 7 méthodes pour chiffrer les messages
- Une technique dérivée est celle de Blaise de Vigenère au XVIème siècle, diplomate au service des rois de France.
  - Il imagine un système de permutation avec une clé, une chaîne de caractères de longueur libre
- Babbage (XIX ème siècle) et la recherche de répétitions
- La méthode Vernam : clés aléatoires aussi longues que les messages, Shannon validera au XIX ème siècle le choix de Vernam
- Les chiffres des Templiers
- Les coincidences de Friedman
- Le carré des Nihilistes (révolution russe)
- Alan Turing et la machine Enigma

Bâtons de Sparte

Cercle de César (58 av JC)



L'écriture secrète des francs-maçons : la clé est la répartition des lettres de l'alphabet dans des "cages"

| L | E | M | A | R | S | O | N | L | A | M | E | I | L | L | E | U | R | E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Q_{17}$ | $U_{21}$ | $E_5$ | $B_2$ | $E_5$ | $C_3$ | $Q_{17}$ | $U_{21}$ | $E_5$ | $B_2$ | $E_5$ | $C_3$ | $Q_{17}$ | $U_{21}$ | $E_5$ | $B_2$ | $E_5$ | $C_3$ | $Q_{17}$ |
| C | Z | R | C | W | V | F | I | Q | C | R | H | Z | G | Q | G | Z | U | V |

| R | E | F | E | R | E | N | C | E | I | N | F | O | R | M | A | T | I | Q | U | E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $U_{21}$ | $E_5$ | $B_2$ | $E_5$ | $C_3$ | $Q_{17}$ | $U_{24}$ | $E_5$ | $B_2$ | $E_5$ | $C_3$ | $Q_{17}$ | $U_{21}$ | $E_5$ | $B_2$ | $E_5$ | $C_3$ | $Q_{17}$ | $U_{21}$ | $E_5$ | $B_2$ |
| M | J | H | J | U | V | I | H | G | N | Q | W | J | W | O | F | W | 2 | L | Z | G |

Les permutations de Vigenere

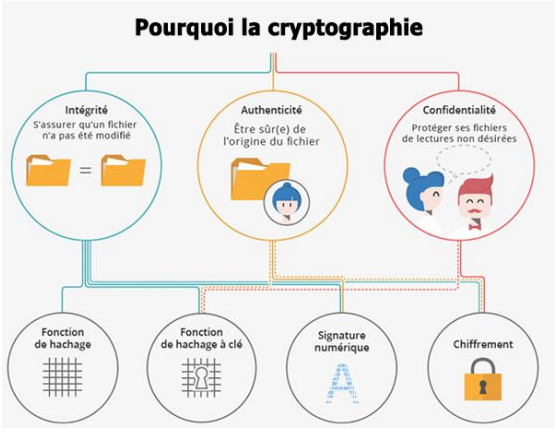
# De quoi parle-t-on, les termes à connaître

- D'abord, un peu de vocabulaire…
- La cryptologie est la science du secret. Elle se distingue en cryptographie, l'écriture secrète et la cryptanalyse (étude des attaques contre les mécanismes de cryptographie).
- La cryptologie déborde la confidentialité des secrets et s'intéresse à l'authenticité des messages (qui ?) et à son intégrité (a-t-il été modifié).
- Pour assurer ses missions, la cryptologie s'appuie sur quatre fonctions : le hashage avec ou sans clé, la signature numérique et le chiffrement.
- La cryptographie a pour objet de développer des méthodes de codage de l'information de sorte qu'il soit difficile de la décoder, si on ne connaît pas la clef qui a permis de le réaliser
- Stéganographie : l'art de cacher une information dans autre chose : image, donnée, vidéo
- Chiffrement : Procédé par lequel on rend la compréhension d'un document impossible à toute personne qui n'a pas la clé de (dé)chiffrement.
- Déchiffrement (quand on connaît la clé) et décryptage (quand on ne la connaît pas)
- On dit chiffrer, pas crypter, ni encrypter.
- En anglais les définitions sont plus floues : "cipher" est l'algorithme utilisé pour chiffrer un message ou une donnée et "encryption" est le processus de conversion de données en se servant de l'algorithme "cypher".




# Le chiffrement où l'art de (se) cacher

## Replacer le chiffrement par rapport aux autres

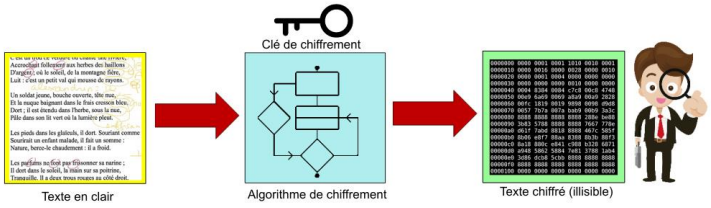


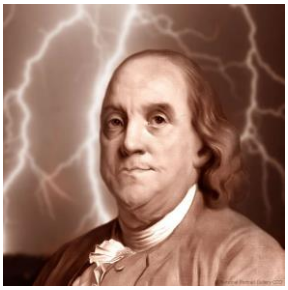

- Un système de chiffrement a pour vocation de cacher "quelque chose"
- La législation a toujours été très rigoureuse, surtout aux Etats-Unis et en France
- Pendant le conflit 39-45, le chiffrement au-delà des clés de 40 bits était puni de…mort
- La législation s'est assouplie…10 ans de prison jusqu'aux années 90
- Le gouvernement avait l'obsession de ne pas pouvoir déchiffrer les messages et documents transmis



# Le chiffrement, mathématiques et clés

- Le chiffrement est fondé sur la transformation mathématique d'un texte à l'aide d'une clé
- L'algorithme peut être réciproque ou symétrique, si pour déchiffrer le texte, on applique la même clé
  - Le "décalage" de César n'est pas symétrique : si la clé est un décalage de 3 lettres dans un sens, il faudrait appliquer un décalage de 3 lettres dans le sens inverse, pour revenir au texte initial…ce qui n'est pas pareil



Benjamin Franklin : Pour qu'un secret soit conservé entre deux personnes, il faut que l'une des deux soit morte…

- Ce qui caractérise un algorithme de chiffrement
  - Le type : symétrique…
  - L'algorithme lui-même et sa complexité
  - La nature de la clé
  - La longueur des clés : plus elle est longue, plus on est protégé contre certaines attaques (force brute, dictionnaires)
  - Le secret des clés, qui elles-mêmes peuvent être chiffrées

# Ne pas confondre clés et hash


-----BEGIN RSA PRIVATE KEY-----
MIICXAIBAAKBgQCqGKukO1De7zhZj6+H0qtjTkVxwTCpvKe4eCZ0FPqri0cb2JZfXJ/DgYSF6vUp
wmJG8wVQZKjeGcjDOL5UlsuusFncCzWBQ7RKNUSesmQRMSGkVb1/3j+skZ6UtW+5u09lHNsj6tQ5
1s1SPrCBkedbNf0Tp0GbMJDyR4e9T04ZZwIDAQABAoGAFijko56+qGyN8M0RVyaRAXz++xTqHBLh
3tx4VgMtrQ+WEgCjhoTwo23KMBAuJGSYnRmoBZM3lMfTKevIkAidPExvYCdm5dYq3XToLkkLv5L2
pIIVOFMDG+KESnAFV7l2c+cnzRMW0+b6f8mR1CJzZuxVLL6Q02fvLi55/mbSYxECQQDeAw6fiIQX
GukBI4eMZZt4nscy2o12KyYner3VpoeE+Np2q+Z3pvAMd/aNzQ/W9WaI+NRfcxUJrmfPwIGm63il
AkEAxCL5HQb2bQr4ByorcMWm/hEP2MZzROV73yF41hPsRC9m66KrheO9HPTJuo3/9s5p+sqGxOlF
L0NDt4SkosjgGwJAFklyR1uZ/wPJjj611cdBcztlPdqoxssQGnh85BzCj/u3WqBpE2vjvyyyvyI5k
X6zk7S0ljKtt2jny2+00VsBerQJBAJGC1Mg5Oydo5NwD6BiROrPxGo2bpTbu/fhrT8ebHkTz2epl
U9VQQSQzY1oZMVX8i1m5WUTLPz2yLJIBQVdXqhMCQBGoiuSoSjafUhV7i1cEGpb88h5NBYZzWXGZ
37sJ5QsW+sJyoNde3xH8vdXhzU7eT82D6X/scw9RZz+/6rCJ4p0=
-----END RSA PRIVATE KEY-----

-----BEGIN PUBLIC KEY-----
MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQCqGKukO1De7zhZj6+H0qtjTkVxwTCpvKe4eCZ0
FPqri0cb2JZfXJ/DgYSF6vUpwmJG8wVQZKjeGcjDOL5UlsuusFncCzWBQ7RKNUSesmQRMSGkVb1/
3j+skZ6UtW+5u09lHNsj6tQ51s1SPrCBkedbNf0Tp0GbMJDyR4e9T04ZZwIDAQAB
-----END PUBLIC KEY-----


- La clé est un artefact utilisé pour chiffrer les documents
- Le hash caractérise (symbolise) un document par une opération mathématique qui lui fait correspondre une valeur binaire sur 128, 192, 256 bits, etc
- MD5 (Message Digest) : 128 bits (32 chiffres hexadécimaux)
- SHA-2 (Secure Hash Algorithm) : 256 ou 512 bits
- SHA-3 : Plus récent (2015) qui utilise les permutations plutôt que le hash, ce qui ajoute de l'imprévisibilité, de la complexité et le rend très résistant

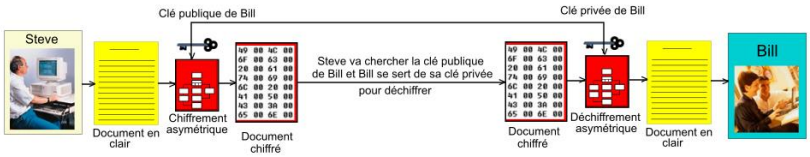
LEMARSONLAMEILLEUREREFERENCEINFORMATIQUE
8459743EF451B7574C8ABE78711D1829F0AB954970479696E0E22A8FEBDF6BF1

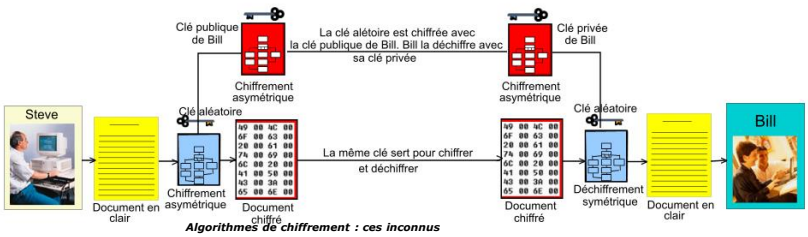
LEMARSON,LAMEILLEUREREFERENCEINFORMATIQUE
6326B40D4673AAA8EF96519801C482DF6A43E48356959A637E9B7E2F81421ABB

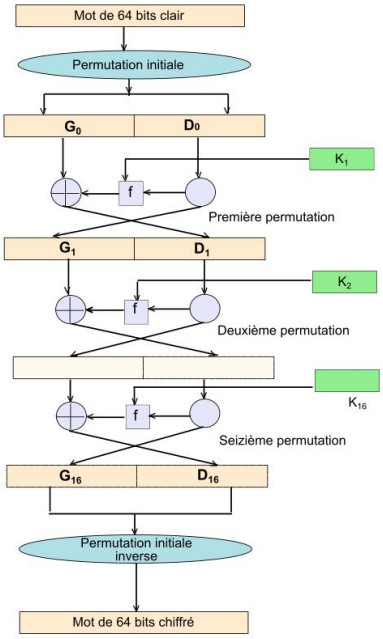
- Une fois le calcul de hash effectué, on ne peut plus revenir en arrière, au document d'origine, contrairement au chiffrement
- Le hash est utilisé en signature électronique, mais pas pour le chiffrement. On peut cependant chiffrer un hash…
- Les deux peuvent être utilisés pour protéger les données, la différence est que l'un est sans pertes et l'autre avec pertes (comme les algorithmes de compression)



# Les principales méthodes




# .e chiffrement symétrique DES

## pour avoir une idée de la mécanique et de la complexité

- Origine : Lucifer d'IBM, approuvé en 1978 par l'ANSI sous le nom de DES (Data Encryption Standard)
- Chiffrement symétrique par blocs de 64 bits, dont 8 bits pour la parité (longueur « utile » de 56 bits).
- L'algorithme consiste à effectuer toute une série d'opérations sur le texte à chiffrer : des combinaisons, des substitutions et des permutations entre le texte à chiffrer et la clé, en faisant en sorte que les opérations puissent se faire dans les deux sens (pour le déchiffrement).
- La clé est codée sur 64 bits et constituée de 16 blocs de 4 bits, notés $K_1$ à $K_{16}$. Etar
- donné que « seuls » 56 bits servent effectivement à chiffrer, il y a $2^{56}$ (soit $7.2*10^{1}$
- clés différentes !
- **L'algorithme DES**
- Fractionnement du texte en blocs de 64 bits (8 octets) ;
- Permutation initiale des blocs : le 58ème bit du bloc de texte de 64 bits se retrouve
- en première position, le 50ème en seconde position et ainsi de suite.
- Découpage des blocs en deux parties : gauche et droite, nommées G et D.
- Etapes de permutation et de substitution répétées 16 fois (appelées **rondes**).
  - $G_1 = D_0$ et $D_1 = G0 + f(D_0, K_1)$
- Recollement des parties gauche et droite pour aboutir à un mot $M_{16} = G_{16}D_{16}$
- puis permutation initiale inverse.



# Le calcul des clés

- Dans une communication de type RSA asymétrique, on utilise deux clés qui sont calculées simultanément
- L'algorithme choisit deux grands nombres premiers P et Q. Plus ils sont grands, plus il est difficile de les casser, avec l'inconvénient de demander plus ressources de traitement et des temps de calculs plus longs
- On calcule N = PQ et Z = (P-1)(Q-1)
- Puis on choisit un nombre E < N qui n'a pas de facteurs communs avec Z, autres que 1 (leur PGCD) : on dit alors que E et Z sont relativement premiers (l'un à l'autre). E est utilisé pour le chiffrement.
- Il faut ensuite trouver un nombre D, tel que ED-1 soit divisible par Z et donc que ED mod Z = 1: D est également utilisé pour le chiffrement
- La clé publique est générée à partir du couple (N,E) et la clé privée à partir des valeurs (N,D)


- Pour casser RSA, très utilisé sur Internet pour authentifier les utilisateurs et garantir la confidentialité bancaire
- La force de RSA tient à ce qu'il ne peut être cassé si l'on n'est pas capable de factoriser des grands nombres
- Pour s'assurer que la clé publique E n'a pas de facteurs communs avec Z, il faut pouvoir le factoriser, ce qui semblait impossible…
- Le GIE (Groupement d'intérêts Economiques) cartes bancaire français avait utilisé le nombre (96 chiffres) :

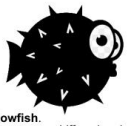
213598703592091008239592270499962879705109534182641740644252416500858395774644508840500943086599 9

"a priori" impossible à factoriser, effectivement résultat du produit de deux facteurs

- Serge Humpich qui a utilisé un logiciel japonais y est parvenu… et à été condamné à 10 mois de prison avec sursis, la décomposition d'une clé publique par factorisation étant un délit
- Ridicule !!!Son tort est peut-être d'avoir voulu négocier
- Ses gains : 10 carnets de tickets de métro (150 € ou 200 $ CAD)
- Depuis, les banques ont augmenté la taille des clés publiques (748 bits actuellement), mais ne sont pas à l'abri d'une attaque

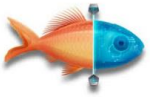
# Panorama des algorithmes utilisés


Véritable succeseur de DES et 3DES. Symétrique, avec clés de 128, 192 ou 256 bits. Créé par Daemen et Rijmen. Utilisé par Blackberry et IBM (Notes). Très performant. Sa principale faiblesse : attaque type "side channel".



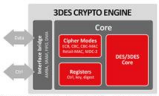
**Blowfish**. Conçu pour chiffrer des données sur des machines 32 bits. Très rapide. Symétrique. Le logiciel BestCrypt l'utilise avec une clé de 448 bits et 16 passages. Sensible aux attaques par dictionnaires.


Symétrique. Chiffrement par blocs, utilisé par PGP. Approuvé par le Gouvernement Canadien. Date de 1996. BestCrypt l'utilise. Sensible aux attaques en texte clair.



**Twofish**. Symétrique avec des clés de 256 bits max. Le NIST (National Institute of Standards and Technology) l'a préconisé pour remplacer DES.


DES (Data Encryption Standard". L'algorithme symétrique le plus connu. A été remplacé par des solutions moins "fragiles". Très rapide, mais il existe de nombreuses solutions pour le "casser". Attaques par force brute.




**3DES.** Dérivé de DES. Symétrique. Utilise 3 clés de 56 bits (168 bits), avec 3 options : clés indépendantes (la plus solide), clés identiques (équivalente à DES) et 2 clés identiques (1 et 3). Vulnérable aux attaques de type "man in the middle" et en "texte clair".


IDEA (International Data Encryption Algorithm). Symétrique. Fondé sur des substitutions et permutations, XOR, Additions et multiplications. Clé de 128 bits. Sensible à la cryptoanalyse dite différentielle et de type "key-schedule"..


RC6 symétrique. Utilise des blocs de 128 bits et des clés variables de 128, 192 et 256 bits, avec 20 phases de traitement. Vulnérabilités connues : attaques par force brute et analytiques.


Asymétrique. Cryptographie à clé publique. La sécurité est basée sur la difficulté de factoriser des grands nombres. Mais avec la puissance des grands nombres...



# Les logiciels de chiffrement













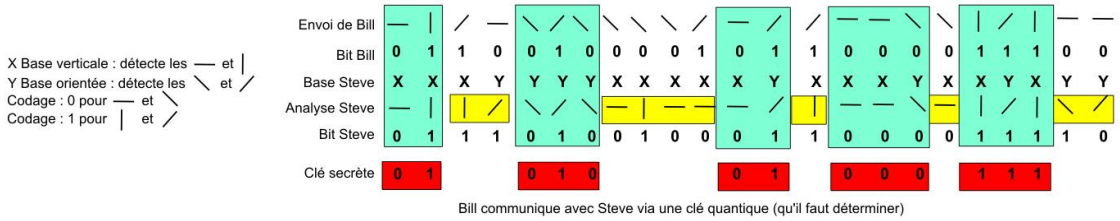

- Plus d'une centaine de solutions sur le marché
- Nombreux gratuits, avec API d'intégration, les autres en souscription
- Chiffrement des équipements liés aux terminaux : disques, SSD…
- Messagerie
- Mots de passe
- Vidéo de bout en bout et outils de collaboration
- Ressources dans le Cloud
- Bases de données et fichiers
- …

# Le chiffrement quantique

- La cryptographie quantique n'est pas de la cryptographie : ce n'est pas le chiffrement d'une information qui serait fondé sur les principes de la mécanique quantique.
- Il faut parler de technique de "distribution de clés", qui permet de distribuer une clé de chiffrement entre deux partenaires, grâce à la mécanique quantique.
- Il existe plusieurs protocoles : BB84 de Bennet et Brassard (1984), qui utilise la polarisation des photons, E91 de Artur Ekert (1991) qui utilise une paire de photons intriqués (effet EPR), mis en évidence par les expérience d'Alain Aspect.

Artur Konrad Ekert FRS est un professeur britannique né en Pologne de physique quantique au Mathematical Institute de l'Université d'Oxford, chercheur en physique quantique et cryptographie

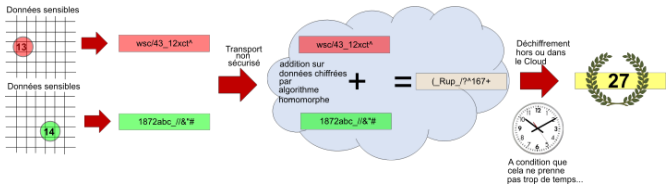
- **Ex de BB84**
- C'est de la magie…
- Fondé sur la polarisation des photons selon 4 directions : H, V, + 45°, +135°
- Le processus consiste à déterminer une clé, connue des deux seuls interlocuteurs : si un pirate veut la lire, il
- La modifie : principe d'Heisenberg
- Pb : le décodage se fait par un analyseur orienté dans l'une des quatre directions, qui ne peut donc détecter qu'un seul plan de polarisation : H ou V (base X) ou +45° ou +135° (Base Y)
  - Si les interlocuteurs utilisent la base X, l'analyseur détectera H ou V, mais produira des résultats aléatoires pour les autres polarisations et même chose pour Y : détection des plans inclinés, mais aléatoire pour les autres
  - Les bases sont dites complémentaires ou incompatibles : il ne peut y en avoir qu'une seule à un instant t (ce serait trop facile…)
  - Il faut donc qu'ils ne retiennent que les bits dont ils sont sûrs, pour déterminer la clé utilisée dans les envois

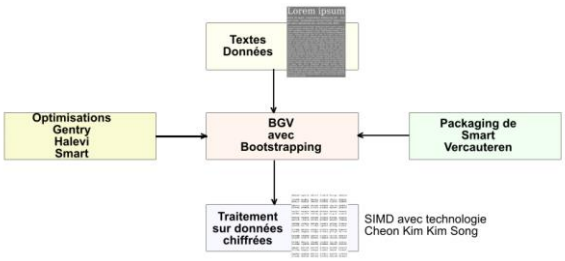



Bill communique avec Steve via une clé quantique (qu'il faut déterminer)



# Le chiffrement homomorphe




- On chiffre les données et le code dans le Cloud et les clés de chiffrement sont conservées chez le client ou le prestataire
- On effectue les traitements directement sur les données chiffrées, sans qu'il soit nécessaire de les décrypter.
- La technique du chiffrement homomorphe n'est pas satisfaisante en termes de temps de traitement. On parle d'un facteur d'un million, voire plus, entre les temps d'exécution portant sur les mêmes données, selon qu'elles sont chiffrées ou non.
- IBM a récemment fait des progrès spectaculaires

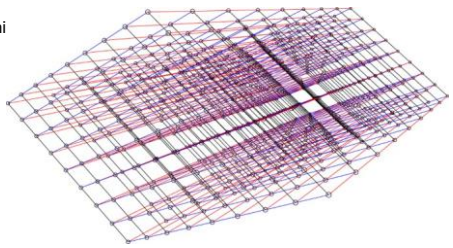
# Inquiétudes pour le ransomware

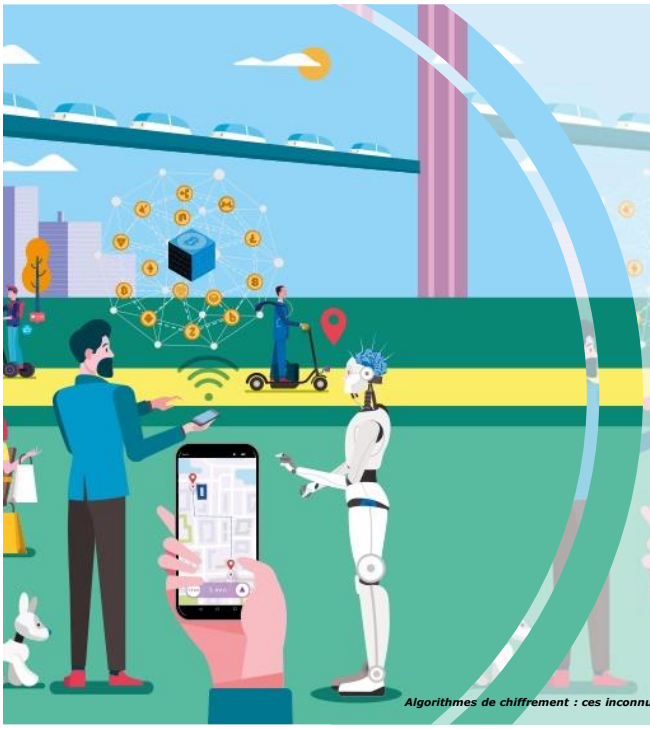
- Les demandes de rançon sont fondées sur le chiffrement des contenus, par des algorithmes à clés secrètes.
- Les criminels créent leurs propres algorithmes, ce qui rend très difficile le décryptage "aveugle".
- Les algorithmes utilisés sont de plus en plus sophistiqués, ce qui est lié aux compétences mathématiques : Russie, Chine, pays de l'ex Europe de l'est.
- Le "ransomware" devrait représenter 20 G$ de pertes globales en 2021, avec la perte de production.
- 1/3 des victimes payent une rançon.
- Stratégie : être paranoïaque, faire des sauvegardes systématiques et… ne pas payer les rançons, sauf cas extrême
- Tendance : les criminels envoient de moins en moins la clé qui a servi à l'attaque.
- Ne jamais laisser accessibles les supports de sauvegarde : NAS…




# Evolutions prévisibles… post quantiques

- Le chiffrement quantique n'est pas exploitable partout : il a besoin d'émetteurs et de récepteurs optiques, de séparateurs de faisceaux… difficiles voire impossibles à mettre dans un smartphone
- Si l'ordinateur quantique "voit le jour", plus aucun algorithme ne sera incassable, il faut donc trouver autre chose
- Une nouvelle génération d'algorithmes va arriver, basée sur des concepts différents, ni quantiques, ni par factorisation
- Fin 2016 : le NIST (National Institute of Standards and Technology) a lancé une consultation pour trouver un algorithme qui résiste aux machines quantiques et aux super-ordinateurs scientifiques
  - 26 algorithmes ont été proposés, classés en 3 familles :
  - Systèmes à treillis, qui existent depuis 1996, des algorithmes décrits dans des espaces mathématiques à N dimensions, 100, 1 000 (devrait aboutir en 2024)
  - Analyse multivariée
  - Dérivé des techniques de codification et détection des erreurs
  - L'algorithme retenu devra être très rapide et fonctionner dans le Web, les smartphones, les capteurs et dispositifs simples, sans que les communications soient ralenties
- Une autre évolution majeure tendra à fournir un meilleur contrôle des clés cryptographiques par les usagers eux-mêmes, avec l'objectif, quand ce sera nécessaire, d'empêcher les fournisseurs et partenaires d'accéder en clair aux données utilisateurs





# Les algorithmes de chiffrement, ces inconnus

30 Octobre 2020

## Nos prochains rendez-vous

| | |
|---|---|
| Vendredi 6 novembre | **: 4G et 5G privés en local, plutôt que Wi-Fi** |
| Vendredi 13 novembre | **: Les certifications pour remplacer les diplômes** |
| Vendredi 20 novembre | **: IA et la démocratie** |
| Vendredi 27 novembre | **: La médecine du futur, les barrières explosent** |
| Vendredi 4 décembre | **: La transformation digitale, mythe ou réalité** |
| Vendredi 18 décembre | **: Panorama des architectures globales du TI** |
| Mercredi 23 décembre | **: Une journée comme les autres en… 2070** |