# Analyse et conception de fonctions de hachage cryptographiques

Thèse présentée et soutenue par

Stéphane Manuel

sous la direction de

Nicolas Sendrier et Daniel Augot

INRIA, Équipe SECRET

23 novembre 2010

# Sommaire

# Sommaire

# Fonctions de hachage

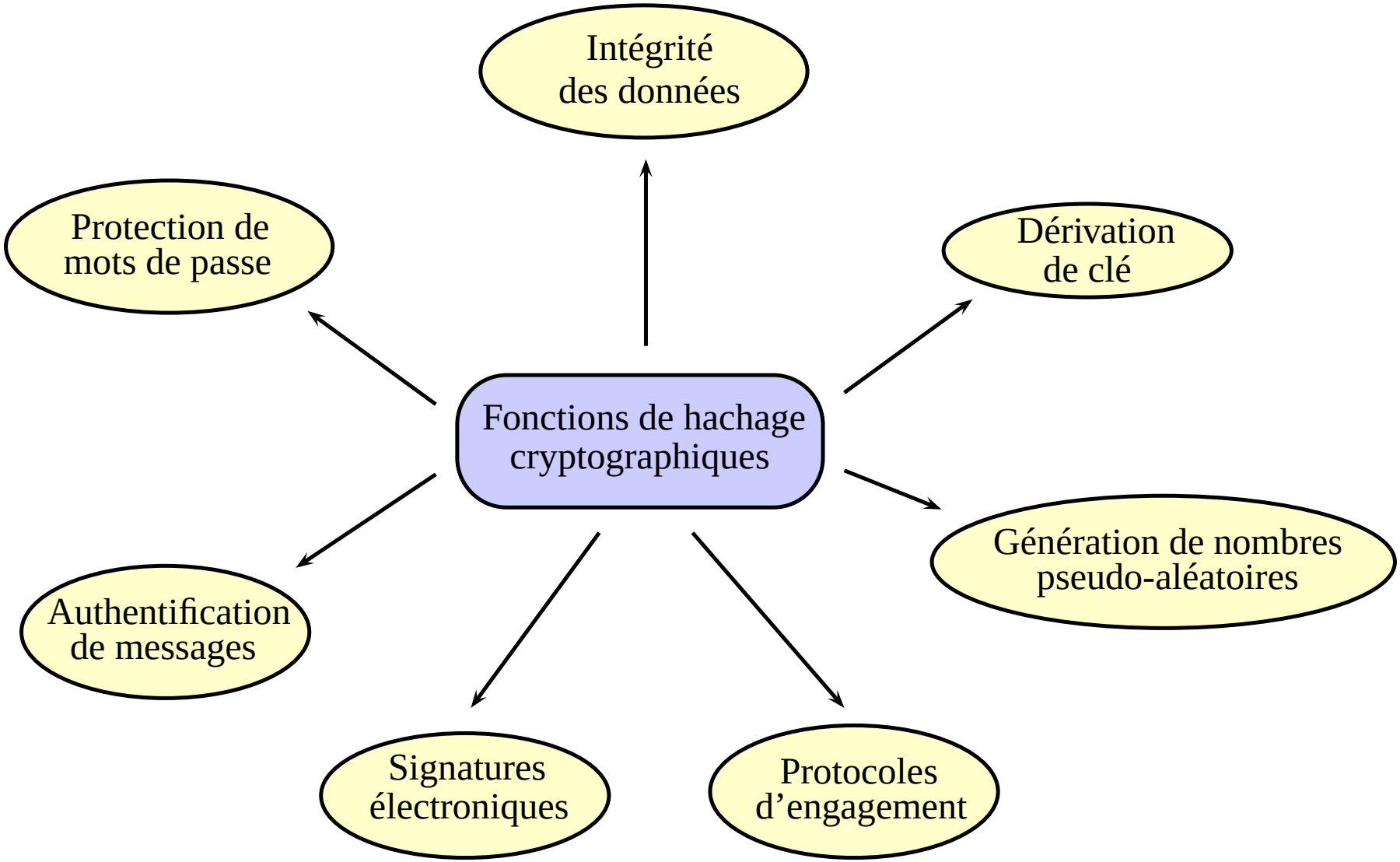
Une fonction de hachage est une fonction qui prend en argument une chaîne de bits de longueur arbitraire finie, et restitue en sortie une chaîne de bits de longueur fixée, nommée empreinte ou haché.

## Définition (Fonction de hachage)

*Une fonction de hachage est une fonction $h$ qui possède les deux propriétés suivantes :*

1. $h : \{0,1\}^* \rightarrow \{0,1\}^n$.
2. *$h$ et $x \in \{0,1\}^*$ donnés, on peut calculer efficacement $h(x)$.*

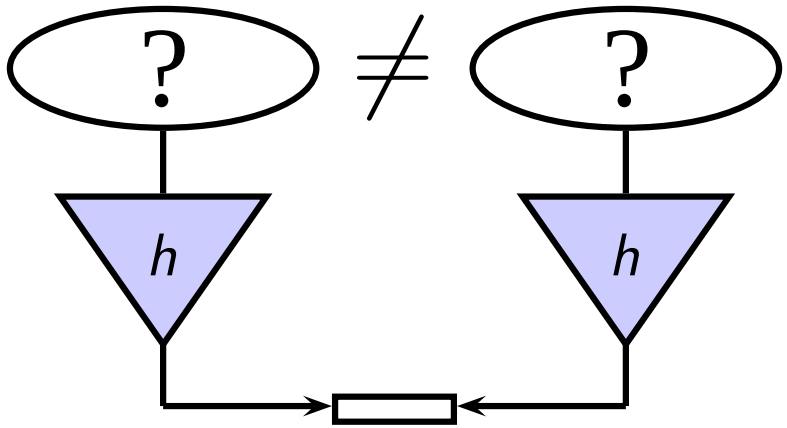
# Domaines d'utilisation en cryptographie



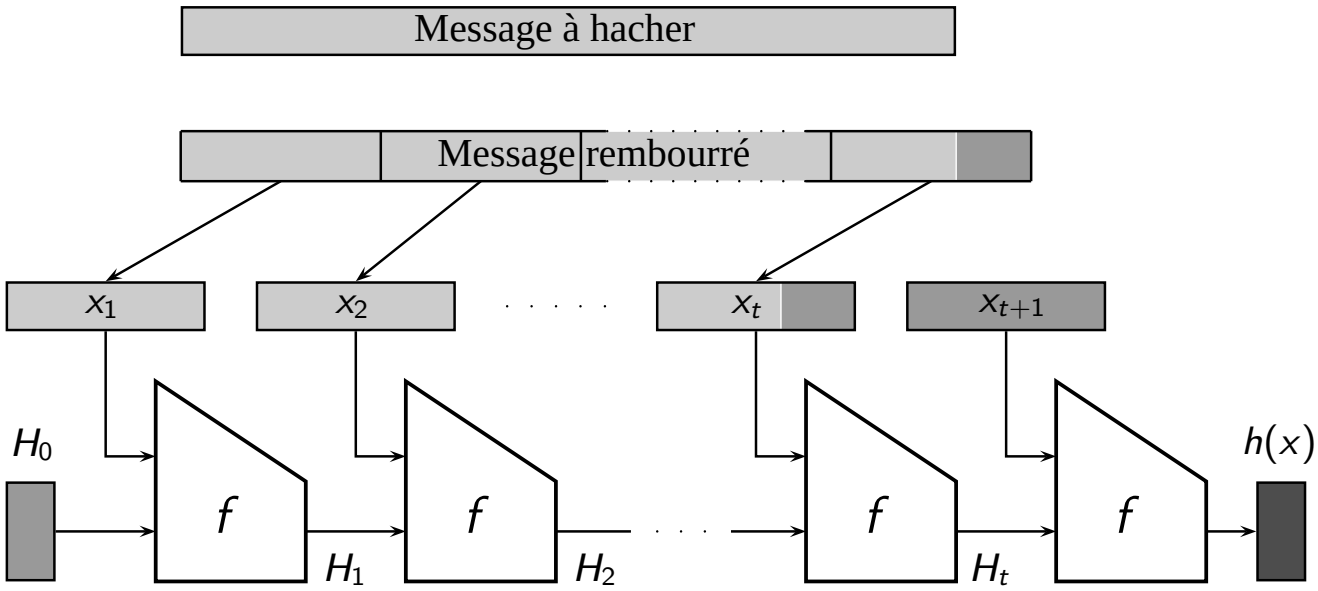
# Résistance aux collisions

## Définition (Résistance aux collisions)

*Trouver un couple de messages* $(x, x') \in \{0,1\}^* \times \{0,1\}^*$, *tel que* $x' \neq x$ *et* $h(x') = h(x)$ *requiert une complexité en* $\mathcal{O}(2^{n/2})$.



# Algorithme de Merkle-Damgård




## Théorème (Conservation de la résistance aux collisions)

*Si la fonction de compression f résiste aux collisions alors la fonction de hachage h résiste aux collisions [Merkle et Damgård 1989].*

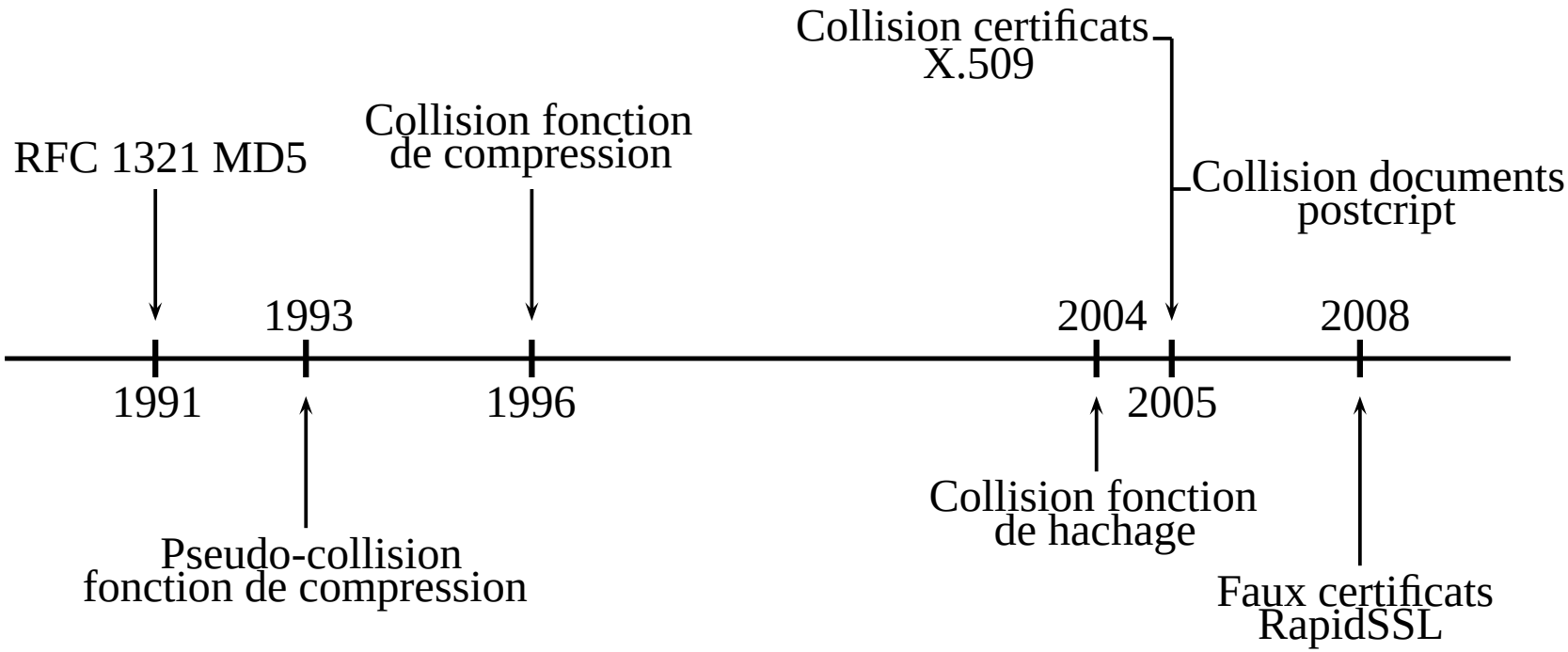
# Famille MD-SHA

Les deux fonctions les plus utilisées en pratique sont les fonctions MD5 et SHA-1.

Ces deux fonctions appartiennent à la famille MD-SHA qui regroupe entre autres les fonctions MD4, MD5, HAVAL, RIPE-MD, SHA-0 et SHA-1.

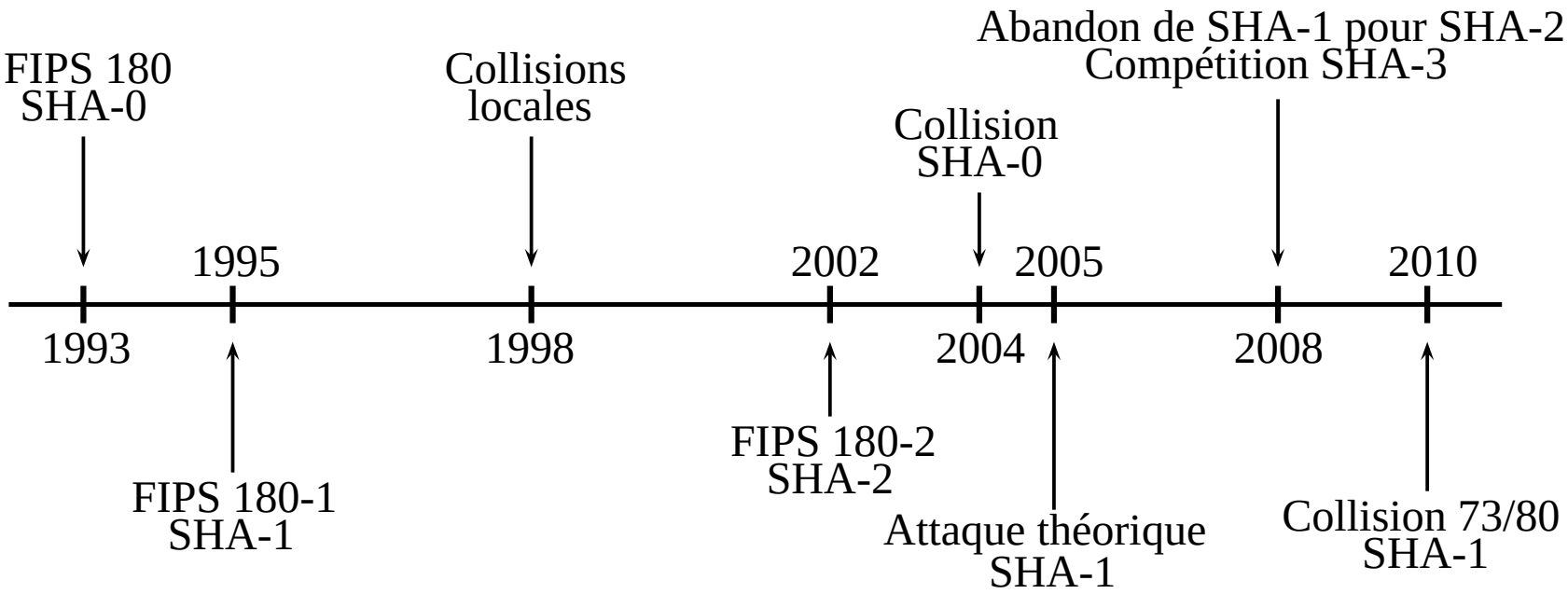
Les fonctions MD5 et SHA-1 sont considérées comme cassées par la communauté des cryptologues.

# Vie et mort de MD5




- 2 ans du standard à la première vulnérabilité
- 11 ans de la première vulnérabilité à la première collision
- 4 ans de la première collision à une attaque dévastatrice

# SHA-0 et SHA-1




- 5 ans du standard à la publication de la première vulnérabilité
- 2010 pas encore de collision
- Faux certificats SSL en 2013 ?

# Mes contributions

Amélioration de l'attaque par collision contre la fonction SHA-0

Étude des caractéristiques linéaires pour la fonction SHA-1

- Classification des vecteurs de perturbations
- Nouveaux résultats sur le comportement des collisions locales

# Sommaire

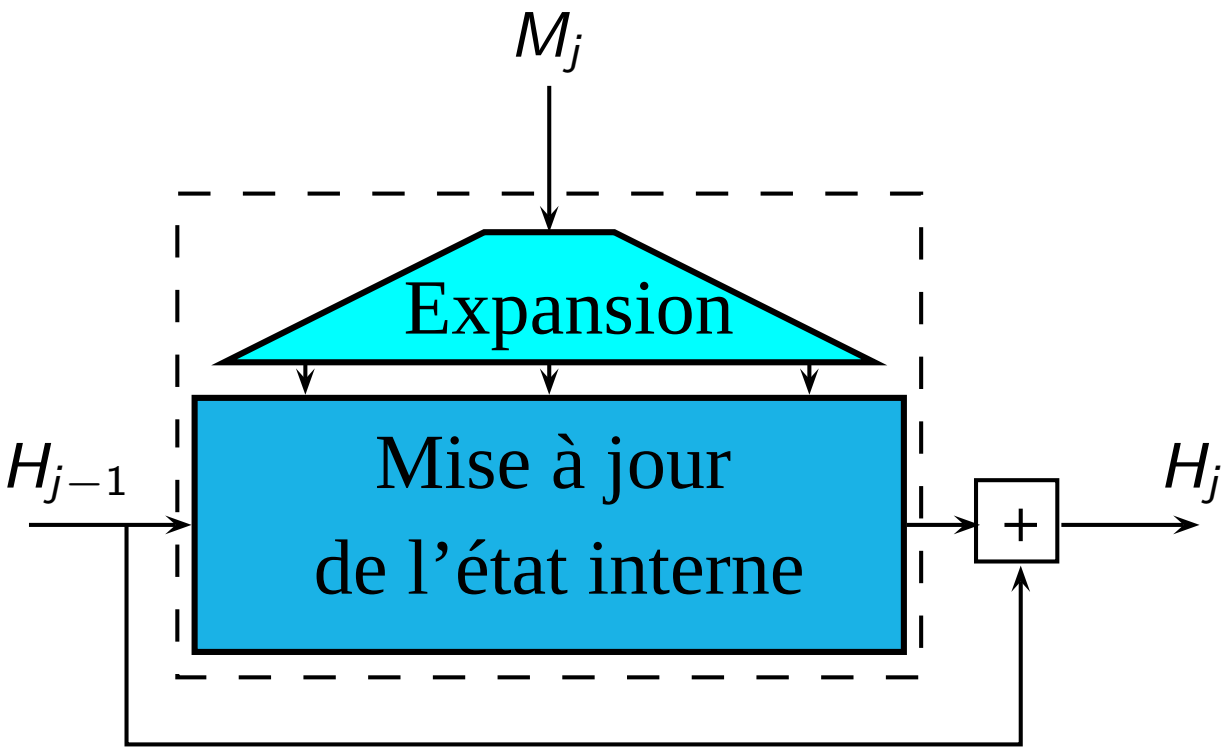
# Fonctions SHA-0 et SHA-1

- Historique
  - 1993 publication du NIST [FIPS 180] : (SHA) SHA-0
  - 1995 nouvelle version [FIPS 180-1] : SHA-1

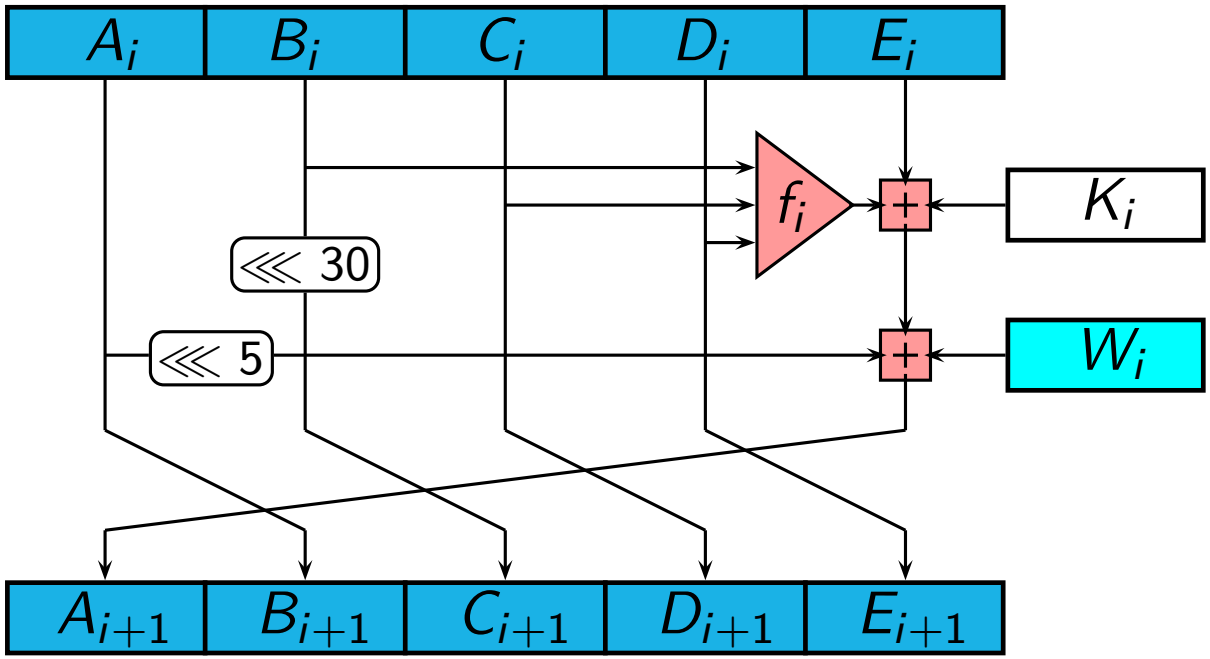
- Description
  - Fonctions de hachage itératives (algorithme de Merkle-Damgård)
  - Entrée : message de longueur strictement inférieure à $2^{64}$ bits
  - Sortie : empreinte de taille 160 bits

# Fonction de compression

- Schéma de chiffrement par bloc ad hoc
- Mode opératoire Davies-Meyer



# Mise à jour de l'état interne



# Mise à jour de l'état interne

- 80 pas :

  Pour $i = 0$ à 79 :

  $A_{i+1} = (A_i \lll 5) + f_i(B_i, C_i, D_i) + E_i + K_i + W_i$

  $B_{i+1} = A_i$

  $C_{i+1} = B_i \lll 30$

  $D_{i+1} = C_i$

  $E_{i+1} = D_i$

- 4 rondes de 20 pas :

| pas $i$ | $f_i(B, C, D)$ | $K_i$ |
|---|---|---|
| $0 \leq i \leq 19$ | $IF = (B \wedge C) \oplus (\overline{B} \wedge D)$ | 5A827999 |
| $20 \leq i \leq 39$ | $XOR = B \oplus C \oplus D$ | 6ED6EBA1 |
| $40 \leq i \leq 59$ | $MAJ = (B \wedge C) \oplus (B \wedge D) \oplus (C \wedge D)$ | 8FABBCDC |
| $60 \leq i \leq 79$ | $XOR = B \oplus C \oplus D$ | CA62C1D6 |

# Expansion du message

Expansion linéaire d'un bloc de message découpé en 16 mots de 32 bits $< M_0, \ldots, M_{15} >$

- SHA-0 :

$$W_i = \begin{cases} M_i & \text{pour } 0 \leq i \leq 15 \\ W_{i-16} \oplus W_{i-14} \oplus W_{i-8} \oplus W_{i-3} & \text{pour } 16 \leq i \leq 79 \end{cases}$$

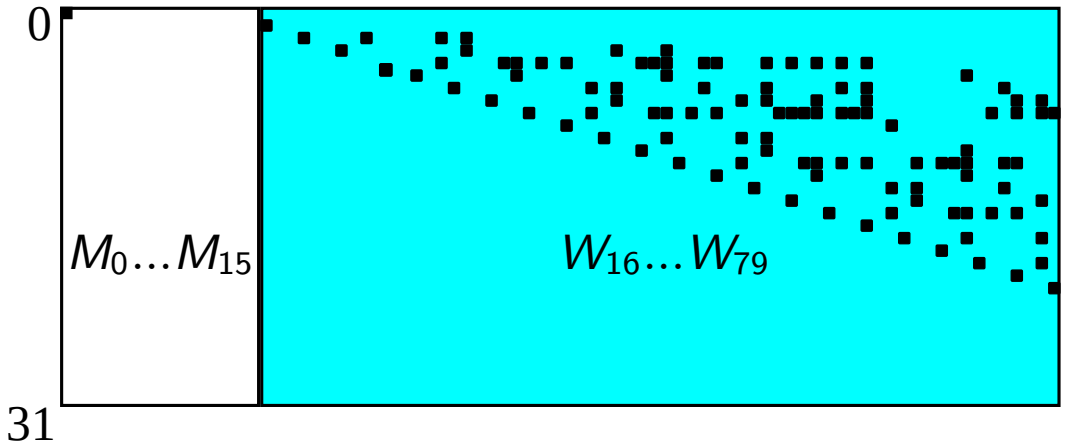
- SHA-1 :

$$W_i = \begin{cases} M_i & \text{pour } 0 \leq i \leq 15 \\ (W_{i-16} \oplus W_{i-14} \oplus W_{i-8} \oplus W_{i-3}) \lll 1 & \text{pour } 16 \leq i \leq 79 \end{cases}$$

# Diffusion

SHA-0

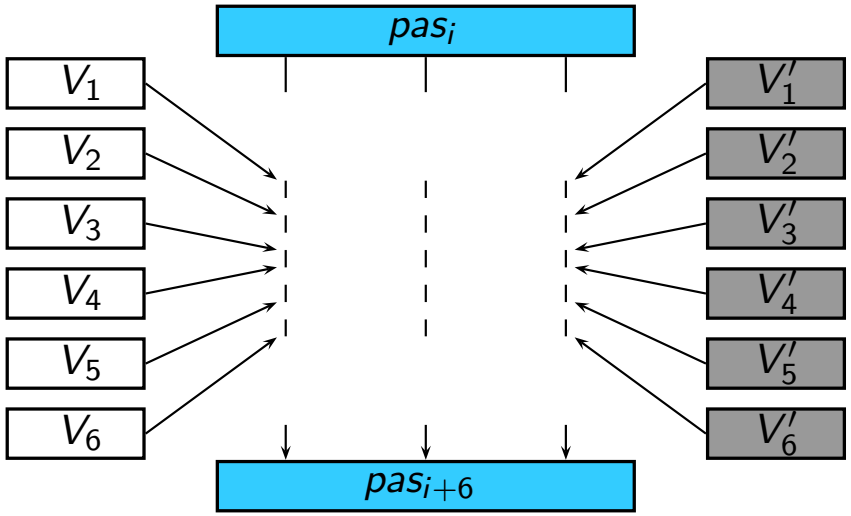

SHA-1

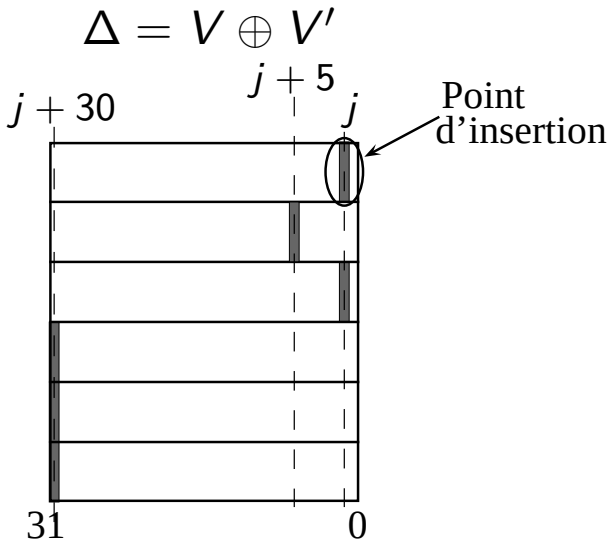


# Collision locale

## Chabaud et Joux [CRYPTO 1998]

- Différentielle sur 6 pas



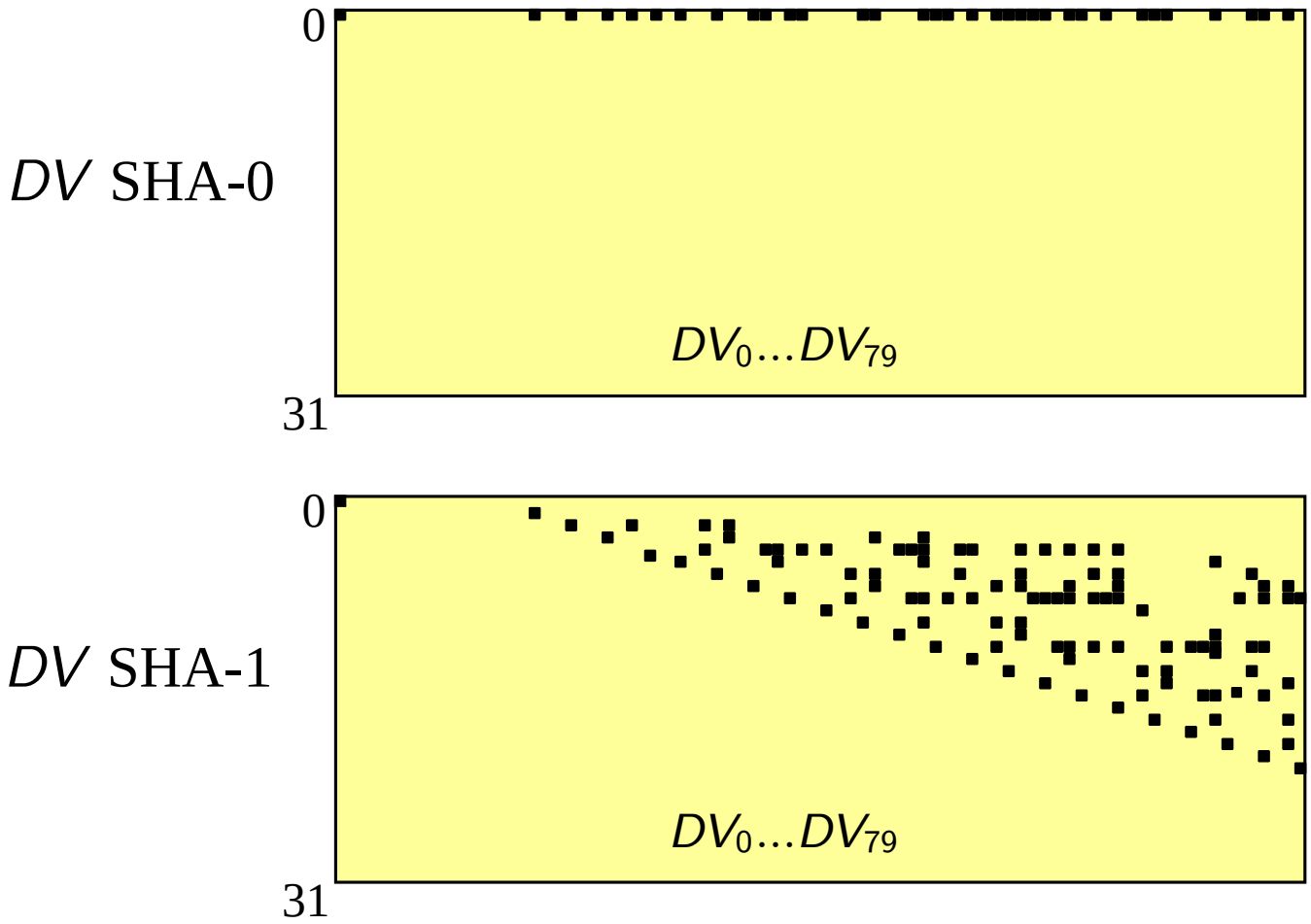

- Probabilités de succès

| $i$ | Position de bit $j$ | | | | f |
|---|---|---|---|---|---|
| | 0 | 1 | 2, . . . , 30 | 31 | |
| $i = 0..19$ | $2^{-5}$ | $2^{-5}$ | $2^{-5}$ | $2^{-4}$ | *IF* |
| $i = 20..39, 60..79$ | $2^{-4}$ | $2^{-2}$ | $2^{-4}$ | $2^{-3}$ | *XOR* |
| $i = 40..59$ | $2^{-4}$ | $2^{-2}$ | $2^{-4}$ | $2^{-4}$ | *MAJ* |

# Vecteur de perturbations

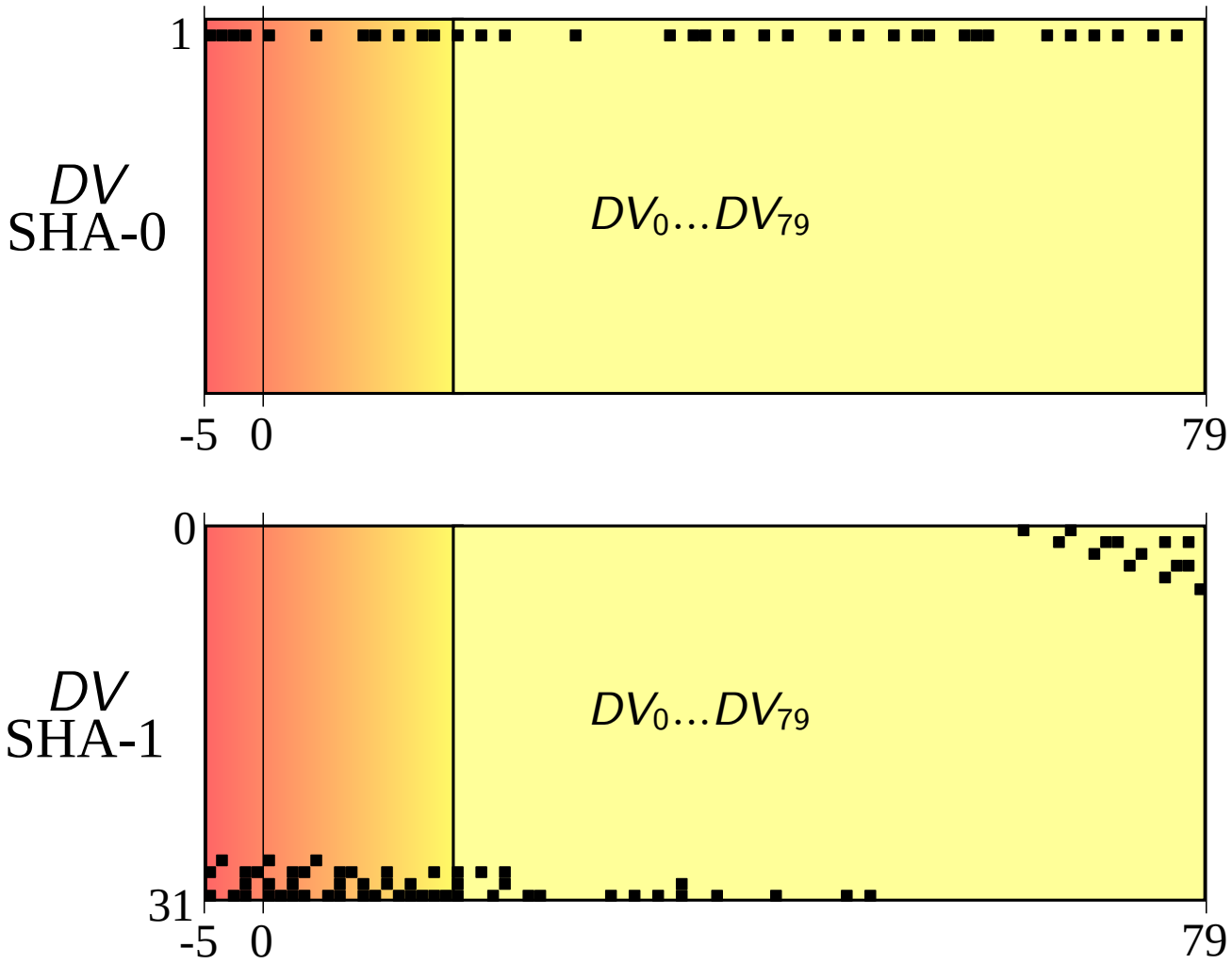
## Chabaud et Joux [CRYPTO 1998]

- Représente les points d'insertion des collisions locales
- Doit être compatible avec l'expansion de message



# Vecteur de perturbations en pratique

- Poids minimisé sur les pas de *MAJ*
- $\approx$ 20 premiers pas traités par la caractéristique non-linéaire



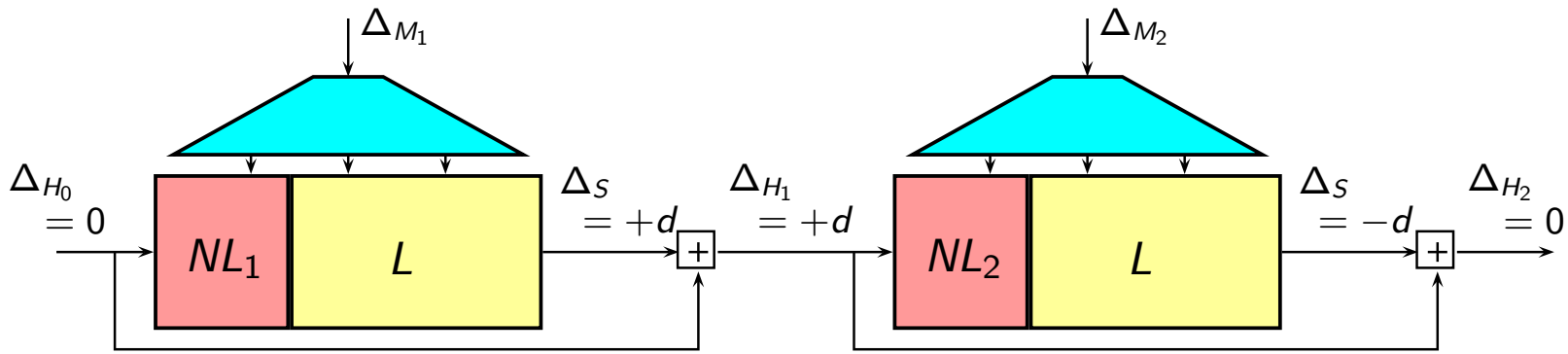
# Attaques par collision contre SHA-0

- Chabaud et Joux [CRYPTO 1998] : modèle des collisions locales
  Première attaque théorique (1 bloc) de complexité $\mathcal{O}(2^{59})$
- Biham *et al.* [EUROCRYPT 2005] : technique des blocs multiples avec utilisation partielle de la non-linéarité de la fonction *IF*
  Première collision (4 blocs) de complexité $\mathcal{O}(2^{49})$
- Wang *et al.* [CRYPTO 2005] : caractéristique non-linéaire et modifications de message
  Collision (2 blocs) de complexité $\mathcal{O}(2^{39})$
- M. et Peyrin [FSE 2008] : technique des blocs multiples, caractéristiques non-linéaires et boomerangs
  Collision (2 blocs) de complexité $\mathcal{O}(2^{33})$

# Principe d'une attaque contre SHA-1

Les principes des attaques par collision contre la fonction SHA-0 s'appliquent à la fonction SHA-1 :

- Collision sur 2 blocs avec la technique des blocs multiples
- Une caractéristique linéaire
- Deux caractéristiques non-linéaires
- Techniques d'accélérations



# Attaques par collision contre SHA-1

- 2005 : Wang *et al.* attaque théorique $\mathcal{O}(2^{69})$ puis $\mathcal{O}(2^{63})$
- 2006 : De Cannière et Rechberger collision pour 64 pas $\mathcal{O}(2^{35})$
- 2007 : De Cannière *et al.* collision pour 70 pas $\mathcal{O}(2^{44})$
- 2007 : Mendel *et al.* collision $\mathcal{O}(2^{60.x})$ calcul abandonné
- 2008 : Peyrin et Joux collision pour 70 pas $\mathcal{O}(2^{39})$
- 2010 : Grechnikov collisions pour 72 $\mathcal{O}(2^{50})$ et 73 pas $\mathcal{O}(2^{52})$

Pas encore de collision pour la fonction SHA-1 :

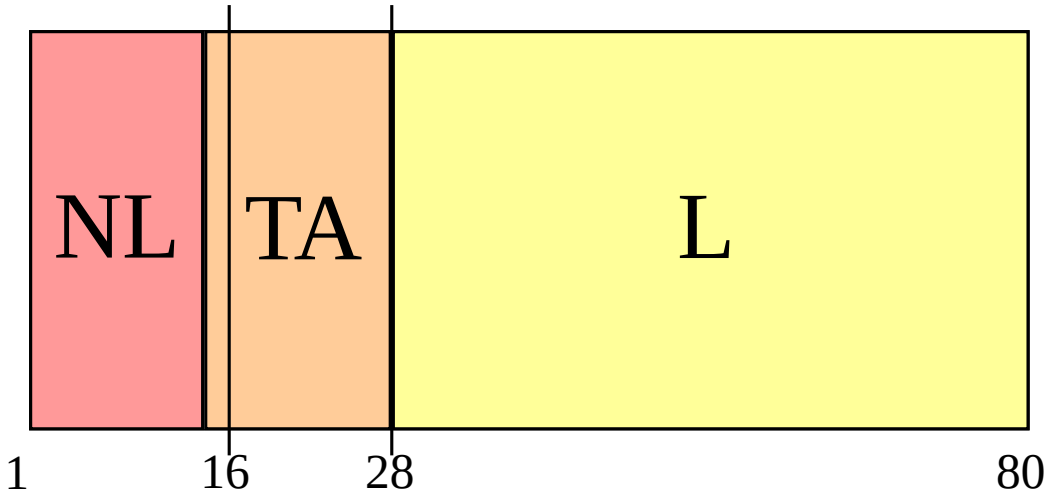
- collision pour 70 pas $\longrightarrow \mathcal{O}(2^{39})$
- collision pour 73 pas $\longrightarrow \mathcal{O}(2^{52})$

# Sommaire

# Complexité des attaques

Aujourd'hui, les techniques de génération de caractéristiques non-linéaires ainsi que les techniques d'accélération sont relativement bien maîtrisées.

Le facteur prépondérant pour la complexité d'une attaque réside dans la caractéristique linéaire.

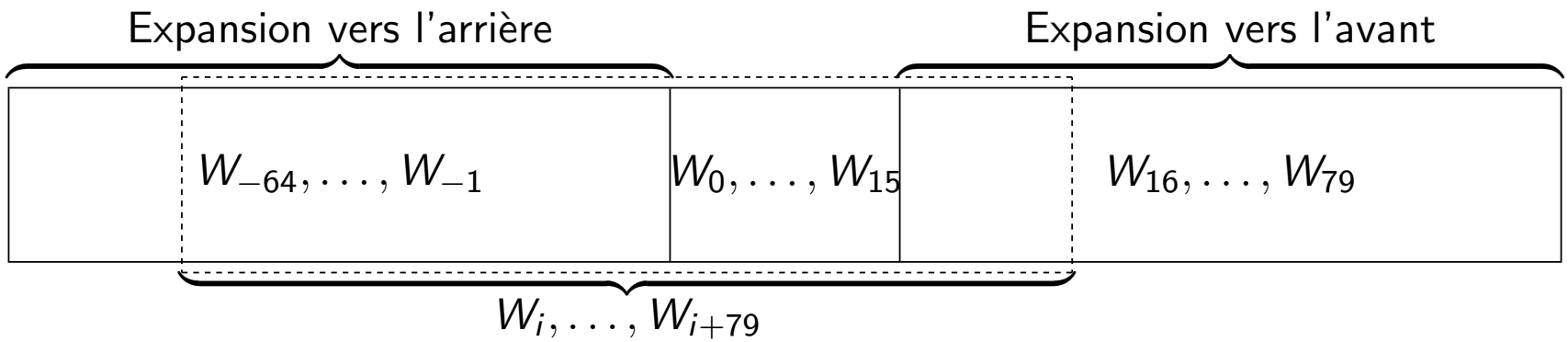



La recherche et l'évaluation des vecteurs de perturbations constituent la piste principale pour améliorer ces attaques.

# Approche mot de code

Fenêtre d'information : 16 mots de 32 bits $(W_0, \ldots, W_{15})$

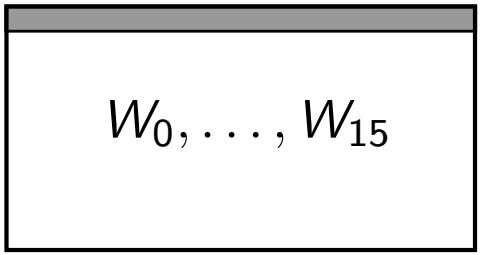
- Deux expansions :
  - Vers l'avant
    $W_i = (W_{i-16} \oplus W_{i-14} \oplus W_{i-8} \oplus W_{i-3}) \lll 1$, pour $16 \leq i \leq 79$
  - Vers l'arrière :
    $W_i = (W_{i+16} \ggg 1) \oplus W_{i+13} \oplus W_{i+8} \oplus W_{i+2}$, pour $-64 \leq i \leq -1$
- Message expansé étendu :



# Compromis sur l'espace de recherche

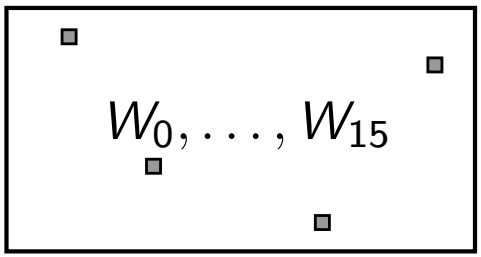
Wang *et al.* [CRYPTO 2005] : *Rectangle Range*

- Poids de la fenêtre d'information $\leq 32$
- Perturbations sur les positions 0 et 1




M. [WCC 2009] : Nouvel algorithme

- Poids de la fenêtre d'information $\leq pw$
- Perturbations sur les positions $0, \ldots, 31$



# Résultats

Recherche exhaustive pour $pw \leq 4$, puis utilisation d'une heuristique pour $pw = 5, 6$

Observations :

- Tous les vecteurs publiés sont obtenus pour $pw \leq 3$
- Similarité de profil des vecteurs les plus performants
- Vecteurs identiques à une permutation circulaire près
- Perturbations concentrées sur les bits de poids fort et de poids faible

Certaines de ces observations étaient déjà présentes dans différents articles sous la forme de remarques.
Cependant aucune explication n'avait été proposée.

# Classification des vecteurs

Relation d'équivalence :

- Invariance par permutation circulaire des bits des mots $(W_0, \ldots, W_{79})$
- Génération à partir du même message expansé étendu

Deux classes pour les vecteurs publiés :

- Type-I : classe contenant le vecteur de Wang *et al.*
- Type-II : classe contenant le vecteur de Jutla et Patthak

Classification :

- Uniformisation de l'ensemble des vecteurs publiés
- Première modélisation et explication des observations présentes dans la littérature

# Nouvelle notation

| Vecteurs de perturbations | Notation |
|---|---|
| Wang *et al.* CRYPTO 2005 | |
| 58 pas | $\mathrm{I}(43, 2)$ |
| 80 pas | $\mathrm{I}(49, 2)$ |
| Rijmen & Oswald CT-RSA 2005 | |
| *Codeword*1 | $\mathrm{I}(45, 1)$ |
| *Codeword*2 | $\mathrm{I}(41, 1)$ |
| *Codeword*3 | $\mathrm{I}(39, 1)$ |
| Jutla & Patthak ePrint 2005 | |
| *Codeword*1 | $\mathrm{I}(52, 0)$ |
| *Codeword*2 | $\mathrm{II}(52, 0)$ |
| *Codeword*3 | $\mathrm{I}(51, 0)$ |
| Pramstaller *et al.* IMA 2005 | $\mathrm{I}(50, 2)$ |
| De Cannière & Rechberger ASIACRYPT 2006 | $\mathrm{I}(35, 2)$ |
| De Cannière *et al.* SAC 2007 | $\mathrm{II}(46, 2)$ |
| Yajima *et al.* ASIACCS 2008 | $\mathrm{II}(56, 2)$ |

# Évaluation des vecteurs de perturbations

Les vecteurs de perturbations représentent une somme de collisions locales avec de possibles entrelacements et il existe différentes approches pour les évaluer.

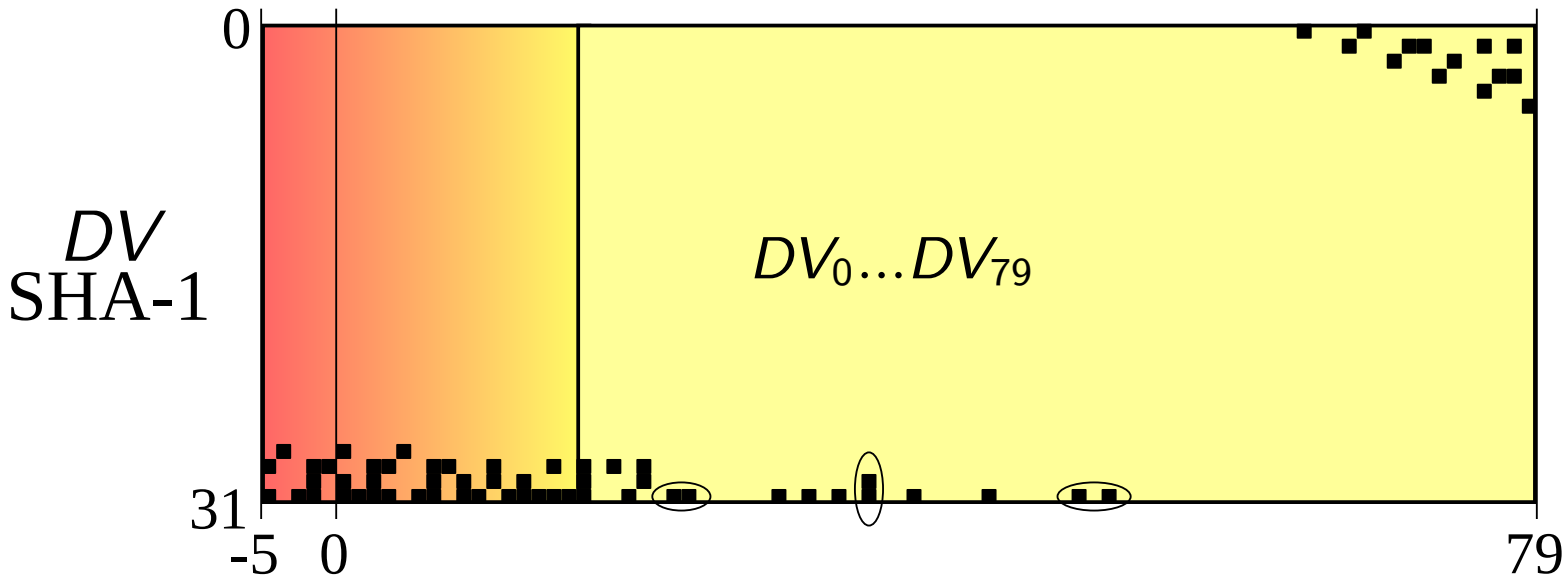
Chacune de ces approches s'appuie sur l'hypothèse que les collisions locales sont indépendantes.
Cependant des cas pathologiques ont été identifiés : le cas des collisions locales consécutives au sein des fonctions *IF* et *MAJ* et la compression de bits pour des collisions locales adjacentes.

La question se pose de vérifier s'il existe d'autres cas pathologiques problématiques ou avantageux.

# Évaluation statistique

Principe :

- Regrouper par petits motifs les entrelacements de collisions locales présents dans les deux classes de vecteurs de perturbations
- Mesurer expérimentalement les probabilités de bon comportement de ces motifs




La première étape consiste à mesurer ces probabilités pour une collision locale isolée.

# Collision locale isolée

| Motif | ----------o |
|---|---|

| Position de bit $b$ | 0 | 1 | 2, . . . , 25 | 26 | 27,28,29 | 30 | 31 |
|---|---|---|---|---|---|---|---|
| Fonction de ronde | Signes | Probabilités mesurées $-\log_2$ | | | | | | |
| IF | + | 4.87 | 5.00 | [4.85, 4.87] | 5.00 | 4.86 | 4.83 | 4.01 |
| | − | 4.87 | 5.00 | [4.85, 4.87] | 5.00 | 4.86 | 4.84 | 4.00 |
| | | (5) | (5) | (5) | (5) | (5) | (5) | (4) |
| XOR | + | 3.68 | 2.00 | [3.90, 3.92] | 4.00 | [3.90, 3.91] | 3.83 | 3.00 |
| | − | 3.68 | 2.00 | [3.90, 3.92] | 4.00 | 3.91 | 3.83 | 3.00 |
| | | (4) | (2) | (4) | (4) | (4) | (4) | (3) |
| MAJ | + | 3.91 | 4.00 | [3.90, 3.92] | 4.00 | [3.90, 3.91] | 3.92 | 4.00 |
| | − | 3.92 | 4.00 | [3.90, 3.92] | 4.00 | 3.91 | 3.91 | 4.00 |
| | | (4) | (4) | (4) | (4) | (4) | (4) | (4) |

*Le motif représente une collion locale en position $j = 0$. Les probabilités théoriques sont notées par des parenthèses. La notation signée indique la direction des collisions locales : "+" de 0 vers 1, "−" de 1 vers 0.*

L'amélioration des probabilités théoriques est due à l'effet de propagation de la retenue.

# Collisions locales adjacentes

| Motif | ----------oo | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Position de bit $b$ | 0 | 1 | 2, . . . , 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Fonction de ronde | Signes | Probabilités mesurées $-\log_2$ | | | | | | | | | |
| XOR | −− | 6.00 | 5.91 | [6.90, 6.91] | 6.96 | 8.00 | 7.90 | 6.91 | 6.87 | 6.36 | 7.00 |
| | ++ | 6.00 | 5.91 | [6.90, 6.90] | 6.96 | 8.00 | 7.91 | 6.90 | 6.87 | 6.36 | 7.00 |
| | | (6) | (6) | (8) | (8) | (8) | (8) | (8) | (8) | (8) | (7) |
| | −+ | 3.68 | 5.91 | [3.90, 3.91] | 3.91 | 3.91 | 7.90 | 3.91 | 3.91 | 3.90 | 3.83 |
| | +− | 3.68 | 5.90 | [3.90, 3.91] | 3.91 | 3.91 | 7.91 | 3.90 | 3.91 | 3.90 | 3.83 |
| | | (4) | (6) | (4) | (4) | (4) | (8) | (4) | (4) | (4) | (4) |
| MAJ | −− | 8.00 | 7.90 | [6.90, 6.91] | 6.96 | 7.99 | 7.90 | 6.91 | 6.91 | 6.95 | 8.00 |
| | ++ | 8.00 | 7.91 | [6.90, 6.91] | 6.96 | 8.00 | 7.91 | 6.91 | 6.91 | 6.95 | 8.00 |
| | | (8) | (8) | (8) | (8) | (8) | (8) | (8) | (8) | (8) | (8) |
| | −+ | 3.91 | 7.91 | [3.90, 3.91] | 3.91 | 3.91 | 7.91 | 3.91 | 3.91 | 3.91 | 3.91 |
| | +− | 3.91 | 7.91 | [3.90, 3.91] | 3.91 | 3.91 | 7.91 | 3.91 | 3.91 | 3.91 | 3.91 |
| | | (4) | (8) | (4) | (4) | (4) | (8) | (4) | (4) | (4) | (4) |

Le principe de la compression de bit fonctionne sauf pour les positions de bit 1 et 26.
On remarque une amélioration inattendue des probabilités théoriques pour des choix de directions a priori inapropriés sur les positions de bit 2..24 et 27..29.

# Collisions locales consécutives

| Motif | ----------o<br>----------o |
|---|---|

| Position de bit $b$ | 0 | 1 | 2, . . . , 24 | 25 | 26 | 27,28,29 | 30 | 31 |
|---|---|---|---|---|---|---|---|---|
| Fonction de ronde | Signes | Probabilités mesurées $-\log_2$ | | | | | | | |
| XOR | $--$ | 5.91 | 4.00 | [5.97, 5.98] | 5.98 | 6.00 | [5.97, 5.98] | 5.91 | 4.00 |
| | $-+$ | 5.91 | 4.00 | [5.97, 5.98] | 5.98 | 6.00 | 5.98 | 5.91 | 4.00 |
| | $+-$ | 5.91 | 4.00 | 5.98 | 5.98 | 6.00 | 5.98 | 5.91 | 4.00 |
| | $++$ | 5.91 | 4.00 | [5.97, 5.98] | 5.98 | 6.00 | [5.97, 5.98] | 5.91 | 4.00 |
| | | (8) | (4) | (8) | (8) | (8) | (8) | (8) | (6) |
| MAJ | $--$ | 10.01 | $> 24$ | [9.88, 9.92] | 9.99 | $> 24$ | [9.91, 9.92] | 9.99 | $> 24$ |
| | $++$ | 10.00 | $> 24$ | [9.88, 9.92] | 10.00 | $> 24$ | [9.90, 9.91] | 10.01 | $> 24$ |
| | | $(\infty)$ | $(\infty)$ | $(\infty)$ | $(\infty)$ | $(\infty)$ | $(\infty)$ | $(\infty)$ | $(\infty)$ |
| | $-+$ | 5.98 | 6.00 | [5.97, 5.98] | 5.98 | 6.00 | 5.98 | 5.98 | 6.00 |
| | $+-$ | 5.98 | 6.00 | [5.97, 5.98] | 5.98 | 6.00 | 5.98 | 5.98 | 6.00 |
| | | (8) | (8) | (8) | (8) | (8) | (8) | (8) | (8) |

*La notation "$> 24$" indique qu'aucune collision ne s'est produite lors de $2^{24}$ essais. La notation "$(\infty)$" désigne un chemin différentiel théoriquement impossible.*

Le cas pathologique de la fonction *MAJ* ne conduit à un chemin différentiel impossible que pour les positions de bit 1, 26 et 31.
La position de bit 31 se révèle dans cette configuration aussi intéressante que la position de bit 1.

# Collisions locales alternées

| Motif | ----------o<br>----------<br>----------o |
|---|---|

| Position de bit $b$ | 0 | 1 | 2, . . . , 24 | 25 | 26 | 27,28,29 | 30 | 31 |
|---|---|---|---|---|---|---|---|---|
| Fonction de ronde | Signes | Probabilités mesurées $-\log_2$ | | | | | | |
| XOR | $--$ | 7.36 | 4.00 | [7.81, 7.82] | 7.82 | 8.00 | [7.81, 7.82] | 7.66 | 6.00 |
| | $-+$ | 7.36 | 4.00 | [7.79, 7.80] | 7.83 | 8.00 | [7.77, 7.81] | 7.66 | 6.00 |
| | $+-$ | 7.35 | 4.00 | [7.79, 7.81] | 7.82 | 8.00 | [7.78, 7.80] | 7.66 | 6.00 |
| | $++$ | 7.36 | 4.00 | [7.81, 7.82] | 7.82 | 8.00 | [7.81, 7.82] | 7.66 | 6.00 |
| | | (8) | (4) | (8) | (8) | (8) | (8) | (8) | (6) |
| MAJ | $--$ | 11.88 | 8.00 | [11.77, 11.82] | 11.91 | $> 24$ | [11.80, 11.82] | 11.91 | $> 24$ |
| | $++$ | 11.91 | 7.99 | [11.79, 11.82] | 11.90 | $> 24$ | [11.81, 11.82] | 11.91 | $> 24$ |
| | $+-$ | 11.89 | $> 24$ | [11.58, 11.62] | 11.90 | $> 24$ | [11.60, 11.63] | 11.93 | 8.00 |
| | $-+$ | 11.92 | $> 24$ | [11.58, 11.63] | 11.90 | $> 24$ | [11.58, 11.63] | 11.89 | 8.00 |
| | | (8) | (8) | (8) | (8) | (8) | (8) | (8) | (8) |

Pour la fonction *MAJ* les probabilités mesurées sont nettement inférieures au probabilités théoriques.
De nouveaux chemins différentiels impossibles apparaissent : sur la position de bit 1 (resp. 31) avec des directions opposées (resp. identiques) et sur la position de bit 26 pour tous les choix de directions.

# Conclusion

Ces évaluations statistiques ont permis de montrer que les probabilités de bon comportement des collisions locales peuvent largement différer des probabilités théoriques.
Elles ont aussi conduit à la mise en évidence de nouveaux cas pathologiques non répertoriés.

Le choix de l'instanciation des directions des collisions locales (uniforme dans la plupart des cryptanalyses) se révèle être d'une importance cruciale pour certaines configurations (motif/position de bit).

Finalement, l'hypothèse d'indépendance des collisions locales ne se vérifie que dans un petit nombre de cas.

# Sommaire

# Cryptanalyse de SHA-0

L'attaque réalisée en collaboration avec Thomas Peyrin et présentée en 2008 à FSE constitue la meilleure attaque publiée contre SHA-0.

Les vecteurs de perturbations sont peu nombreux ($2^{16} - 1$) et leur évaluation est simple.
Les meilleurs vecteurs sont bien identifiés.

La meilleure piste de recherche pour une amélioration de l'attaque réside probablement dans les techniques d'accélération.

# Cryptanalyse de SHA-1

La classification proposée fournit un modèle qui unifie l'ensemble des vecteurs publiés et explique les remarques et observations présentes dans la littérature.

Les mesures expérimentales montrent que l'hypothèse commune sur l'indépendance des collisions locales est fausse pour des motifs présents dans les vecteurs de perturbations publiés.
L'instanciation des directions des collisions locales devient un paramètre non trivial de l'attaque.

La poursuite de l'étude du comportement des entrelacements de collisions locales constitue une bonne piste de recherche pour l'obtention d'une collision pour SHA-1.

# Autres travaux de recherche

Conception de fonctions de hachage :

- Proposition de la fonction XOR-HASH
  (en collaboration avec Nicolas Sendrier WEWoRC 2007)
- Contribution au processus de développement de la fonction FSB
  (Daniel Augot, Matthieu Finiasz, Philippe Gaborit et Nicolas Sendrier 2008)

Proposition d'algorithmes de colorisation d'images segmentées
(en collaboration avec Catherine Sauvaget, Jean-Noel Vittaut, Jordane Suarez et Vincent Boyer SITIS 2010)

# Vie et mort de MD5

- 1991 RFC 1321 : Rivest
- 1993 pseudo-collision fonction de compression : Den Boer et Bosselars
- 1996 collision fonction de compression : Dobbertin
- 2004 collision fonction de hachage : Wang *et al.*
- 2005 collision certificats X.509 : Lenstra, Wang et de Weger
- 2005 collision documents postscript : Daum et Lucks
- 2008 faux certificats RapidSSL : Sotirov *et al.*

# Collision locale

## Chabaud et Joux [CRYPTO 1998]

Modèle des collisions locales

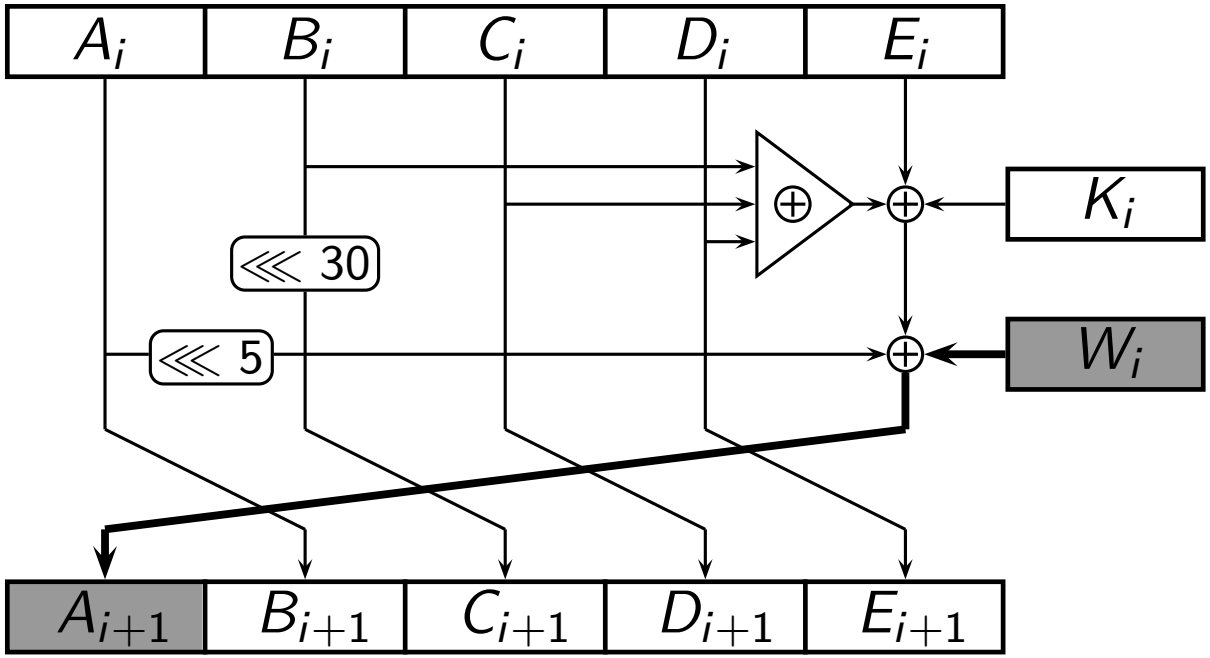
- Introduire une perturbation sur une position de bit $j$ $(0 \leq j \leq 31)$
- Compenser l'impact sur l'état interne par 5 corrections

$$
\begin{aligned}
A_{i+1} &= (A_i \lll 5) + f_i(A_{i-1}, A_{i-2} \ggg 2, A_{i-3} \ggg 2) + A_{i-4} \ggg 2 + K_i + \mathbf{W_i^j} \\
A_{i+2} &= (A_{i+1} \lll 5) + f_i(\mathbf{A_i^{j+5}}, A_{i-1} \ggg 2, A_{i-2} \ggg 2) + A_{i-3} \ggg 2 + K_{i+1} + \mathbf{W_{i+1}^{j+5}} \\
A_{i+3} &= (A_{i+2} \lll 5) + f_i(A_{i+1}, \mathbf{A_i^j} \ggg 2, A_{i-1} \ggg 2) + A_{i-2} \ggg 2 + K_{i+2} + \mathbf{W_{i+2}^j} \\
A_{i+4} &= (A_{i+3} \lll 5) + f_i(A_{i+2}, A_{i+1} \ggg 2, \mathbf{A_i^{j+30}} \ggg 2) + A_{i-1} \ggg 2 + K_{i+3} + \mathbf{W_{i+3}^{j+30}} \\
A_{i+5} &= (A_{i+4} \lll 5) + f_i(A_{i+3}, A_{i+2} \ggg 2, A_{i+1} \ggg 2) + \mathbf{A_i^{j+30}} \ggg 2 + K_{i+4} + \mathbf{W_{i+4}^{j+30}} \\
A_{i+6} &= (A_{i+5} \lll 5) + f_i(A_{i+4}, A_{i+3} \ggg 2, A_{i+2} \ggg 2) + \mathbf{A_{i+1}^{j+30}} \ggg 2 + K_{i+5} + \mathbf{W_{i+5}^{j+30}}
\end{aligned}
$$

# Collision locale

## Étape 1

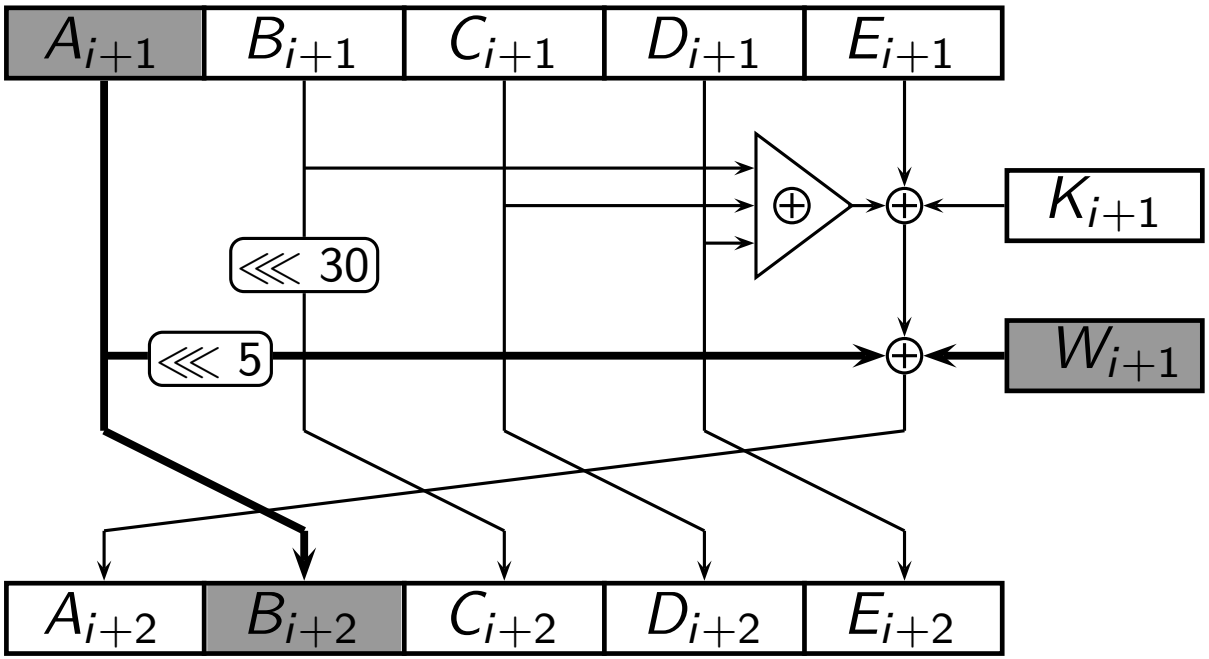
Approximation linéaire de la fonction de mise à jour des registres




$$\overset{j}{A}_{i+1} = (A_i \lll 5) \oplus B_i \oplus C_i \oplus D_i \oplus E_i \oplus \overset{j}{W}_i \oplus K_i$$

# Collision locale

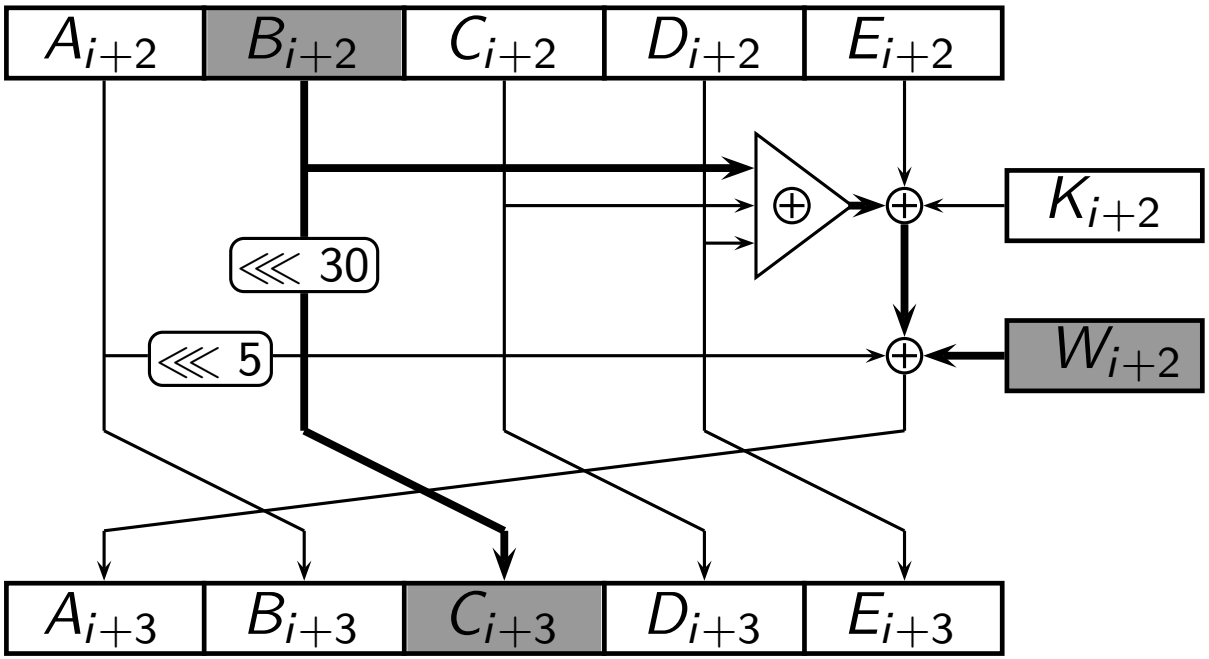
## Étape 2




$$A_{i+2} = (\overset{j}{A_{i+1}} \lll 5) \oplus B_{i+1} \oplus C_{i+1} \oplus D_{i+1} \oplus E_{i+1} \oplus \overset{j+5}{W_{i+1}} \oplus K_{i+1}$$

# Collision locale

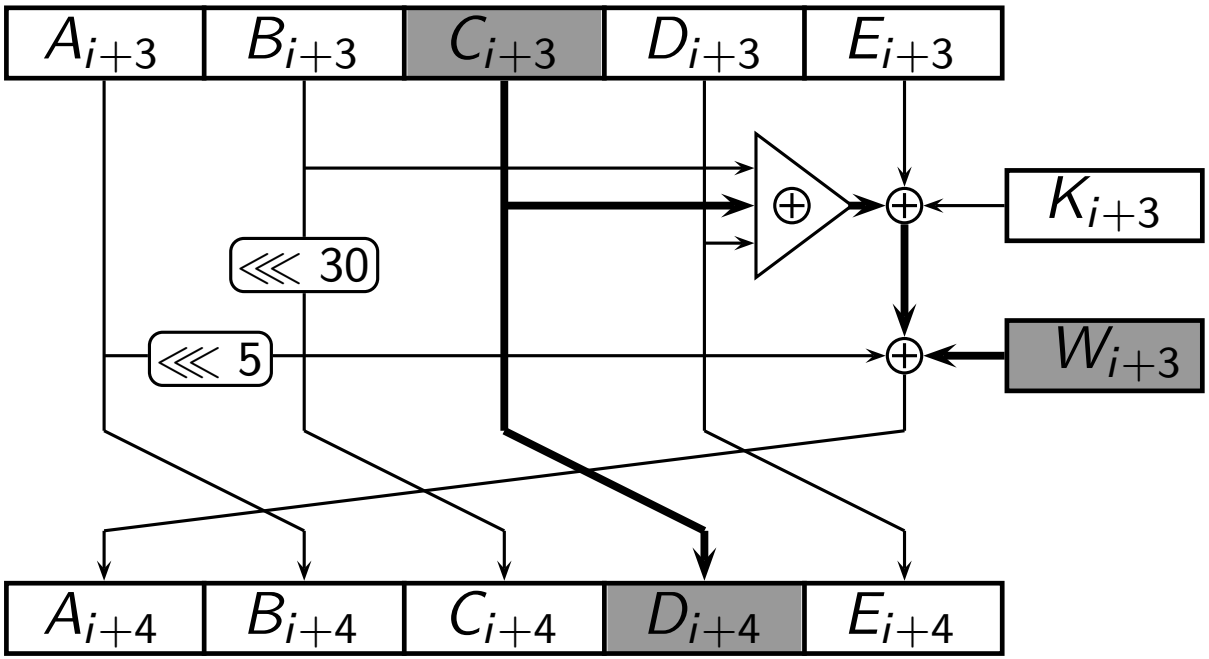
## Étape 3




$$A_{i+3} = (A_{i+2} \lll 5) \oplus \overset{j}{B}_{i+2} \oplus C_{i+2} \oplus D_{i+2} \oplus E_{i+2} \oplus \overset{j}{W}_{i+2} \oplus K_{i+2}$$

# Collision locale

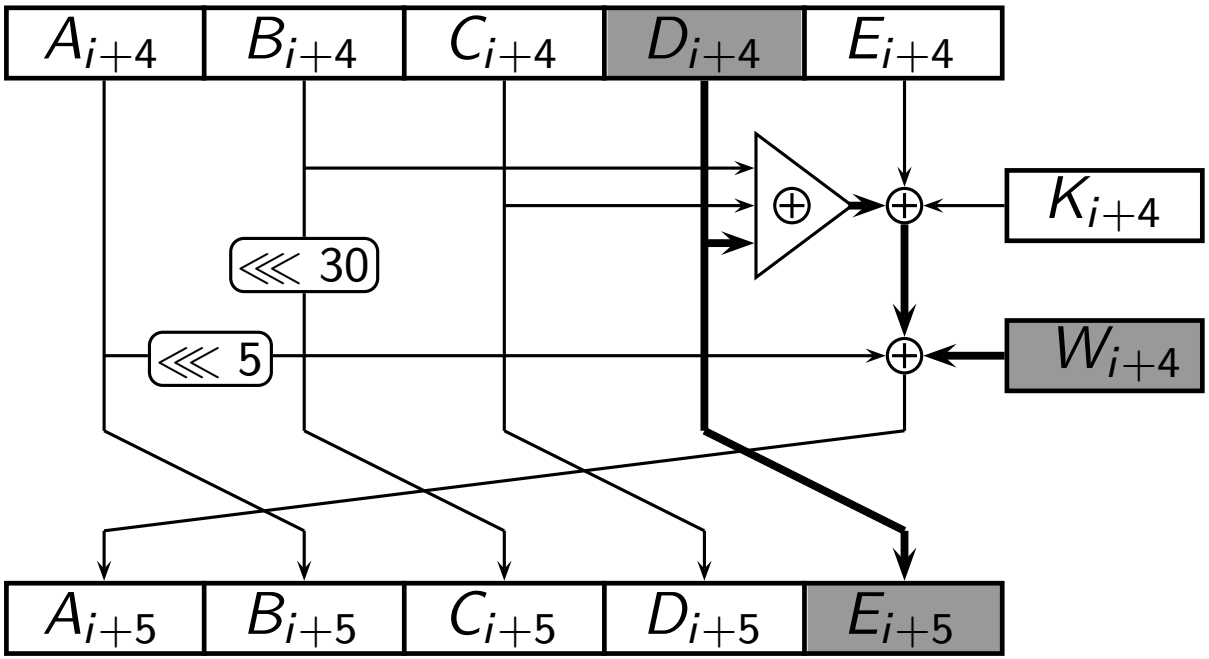
## Étape 4




$$A_{i+4} = (A_{i+3} \lll 5) \oplus B_{i+3} \oplus \overset{j+30}{C_{i+3}} \oplus D_{i+3} \oplus E_{i+3} \oplus \overset{j+30}{W_{i+3}} \oplus K_{i+3}$$

# Collision locale

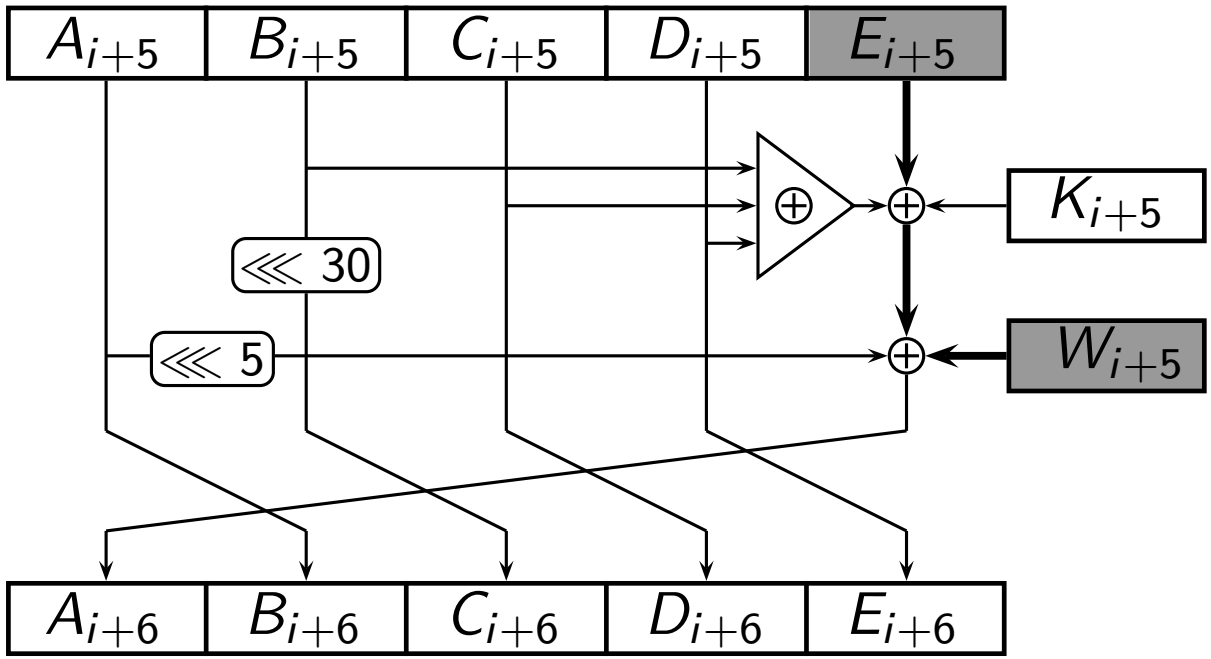
## Étape 5




$$A_{i+5} = (A_{i+4} \lll 5) \oplus B_{i+4} \oplus C_{i+4} \oplus \overset{j+30}{D_{i+4}} \oplus E_{i+4} \oplus \overset{j+30}{W_{i+4}} \oplus K_{i+4}$$

# Collision locale

## Étape 6




$$A_{i+6} = (A_{i+5} \lll 5) \oplus B_{i+5} \oplus C_{i+5} \oplus D_{i+5} \oplus \overset{j+30}{E_{i+5}} \oplus \overset{j+30}{W_{i+5}} \oplus K_{i+5}$$

# Premier bloc

Phase de précalcul :

- Trouver une bonne caractéristique linéaire
- Instancier les directions des collisions locales
  - Fixer le signe de la différence de sortie
- Traduire l'instanciation en équations sur les bits de message
- Ajouter des boomerangs
  - Équations supplémentaires
- Construire une première caractéristique non-linéaire
  - Compatible avec toutes les équations

Phase de recherche :

- Trouver une paire de messages vérifiant la différence de sortie
  - Utilisation des modifications de message et des boomerangs

# Second Bloc

Phase de précalcul :

- Instancier les directions des collisions locales
  - Fixer le signe opposé pour la différence de sortie
- Traduire l'instanciation en équations sur les bits de message
- Ajouter des boomerangs
  - Équations supplémentaires
- Construire une seconde caractéristique non-linéaire
  - Compatible avec toutes les équations et la différence en entrée

Phase de recherche :

- Trouver une paire de messages vérifiant la différence de sortie avec un signe opposé
  - utilisation des modifications de message et des boomerangs

# Mendel *et al.* [FSE 2006]

Collision locale isolée

- Fonction *XOR* :

$$P(j) = \begin{cases} 2^{-4} + 2^{-6} & \text{pour } j = 0 \\ 2^{-2} & \text{pour } j = 1 \\ \sum_{k=1}^{27-j} 2^{-4k} & \text{pour } j = 2, \ldots, 26 \\ 2.2^{-4.(32-j)} + \sum_{k=1}^{31-j} 2^{-4k} & \text{pour } j = 27, \ldots, 31 \end{cases}$$

- Fonction *MAJ* :

$$P(j) = \begin{cases} 2^{-4} + 2^{-8} & \text{pour } j = 0 \\ 2^{-4} & \text{pour } j = 1 \\ \sum_{k=1}^{27-j} 2^{-4k} & \text{pour } j = 2, \ldots, 26 \\ \sum_{k=1}^{32-j} 2^{-4k} & \text{pour } j = 27, \ldots, 30 \\ 2^{-4} & \text{pour } j = 31 \end{cases}$$

# Mendel *et al.* [FSE 2006]

Deux collisions locales adjacentes, fonction *XOR* :

$$P(j,j+1) = \begin{cases} 2^{-4} + 2^{-6} & \text{pour } j = 0 \\ 2^{-4} + 2^{-8} & \text{pour } j = 1 \\ \sum_{k=1}^{27-j} 2^{-4k} & \text{pour } j = 2, \ldots, 25 \\ 2^{-4} + 2^{-8} & \text{pour } j = 26 \\ 2.2^{-4.(32-j)} + \sum_{k=1}^{31-j} 2^{-4k} & \text{pour } j = 27, \ldots, 30 \end{cases}$$