TECHNIQUES
DE L'INGÉNIEUR

Réf. : **E3440 V2**

Date de publication :
10 décembre 2019

Cartes à puces - Technologie et cybersécurité

Cet article est issu de : **Électronique - Photonique | Électronique**

par **Jean-Pierre TUAL, Stéphane GRELLIER, Joseph LEIBENGUTH, Philippe PROUST**

Mots-clés
sécurité numérique | Internet des Objets | cybersécurité | cartes à puce | éléments de sécurité | passeports électroniques

Résumé La carte à puces désigne les supports de sécurité contenant un circuit électronique intégré capable de mémoriser ou de traiter les informations. La carte à puces est à la base de la sécurité de nombreux systèmes informatiques. Elle a fait ses preuves dans plusieurs secteurs en tant que moyen de paiement, d'identification ou d'authentification sûre pour les utilisateurs. Cet article traite des technologies semi-conducteur pour les cartes à puces, de l'importance de la cryptologie et de la sécurité physique et logique. Il décrit aussi les différents types de cartes à puces, leur architecture, leur construction et fabrication, et s'intéresse aux systèmes d'exploitation avant de présenter les perspectives d'avenir.

Keywords
digital security | Internet of Things | cybersecurity | smart cards | secure elements | electronic passports

Abstract The expression smart cards refers to security supports containing an electronic circuit capable of memorizing or processing information. The smart card is at the basis of the security of several IT systems. It has proved its worth in many sectors as a secure mean of payment, identification or authentication for end-users. This article discusses semiconductor technologies for smart cards, the importance of cryptology, and physical and logical security. It also describes the different types of smart cards, their architecture, construction and manufacturing, and looks at operating systems before presenting the future.

Pour toute question :
Service Relation clientèle
Techniques de l'Ingénieur
Immeuble Pleyad 1
39, boulevard Ornano
93288 Saint-Denis Cedex

Par mail :
infos.clients@teching.com
Par téléphone :
00 33 (0)1 53 35 20 20

Cartes à puces

Technologie et cybersécurité

par **Jean-Pierre TUAL**
Ancien directeur des relations industrielles,
Direction technologie et innovation, Gemalto
Auteur de la version originale de l'article 2007

Stéphane GRELLIER
Mobile software security & services manager ,
Gemalto, Meudon, France
Auteur de la version actualisée de 2019

Joseph LEIBENGUTH
Physical document security R&D product director – Technical advisor,
Gemalto, Saint-Cloud, France
Auteur de la version actualisée de 2019

et **Philippe PROUST**
Embedded & core security director,
Gemalto, Géménos, France
Auteur de la version actualisée de 2019

*Le nom de **carte à puces** est couramment utilisé pour désigner des supports de sécurité en matière plastique aux mêmes dimensions qu'une carte de crédit et qui contiennent un circuit électronique intégré capable de mémoriser ou de traiter les informations. L'AFNOR (Association Française de Normalisation) a retenu le terme de **cartes à microcircuits à contacts**, car l'interface électrique de ces cartes est assurée par des liaisons galvaniques. Des cartes à interface sans contact, basée sur la liaison radiophonique, se sont imposées depuis plusieurs années, et ont permis l'adoption de nouveaux facteurs de forme comme le passeport électronique. Ils sont aujourd'hui au cœur de la croissance avec l'adoption du paiement sans contact par un nombre croissant de pays.*

*La carte à puces, dont la gestation a pu sembler très longue, est à la base de la sécurité des systèmes informatiques. Elle a désormais fait ses preuves dans de nombreux secteurs de l'activité humaine en tant que moyen de paiement, d'identification sur les réseaux fixes (de type Internet), mobiles (GSM ou UMTS) ou multimédia (télévision à péage), d'authentification pour les services gouvernementaux (cartes d'identité, passeports électroniques). La **carte SIM, ou USIM**, clé d'accès aux réseaux de téléphonie mobile, et son équivalent **Secure Element** (SE) pour l'internet des objets (IoT), au facteur de forme plus petit, constitue probablement le composant électronique intelligent le plus utilisé dans le monde (5,6 milliards d'unités vendues en 2017 !). De même, la **carte bancaire** à microcalculateur, dont l'utilisation s'est généralisée en France depuis 1992, a connu une croissance quasi exponentielle avec une généralisation de son utilisation en Europe, au Japon, en Chine, ainsi qu'aux États-Unis en version sans contact.*

Grâce aux progrès continuels des semi-conducteurs, des technologies de fabrication et de l'évolution des techniques de programmation utilisables, des développements considérables de la carte à puces ont pu avoir lieu et se poursuivent. La carte à puces et ses variantes constituent, pour beaucoup d'applications, une solution particulièrement bien adaptée aux enjeux socio-économiques de notre société.

L'objet de cet article est d'apporter une vue d'ensemble sur les briques technologiques développées spécifiquement pour les cartes à puces et sur leur importance dans la fiabilité et la sécurité physique et logique de ce produit. La diversité des compétences requises pour concevoir les cartes à puces, produire le composant électronique et la carte dans son ensemble, fabriquer les cartes à plusieurs milliards d'unités par an, explique la force de cette industrie et le potentiel qu'elle offre dans le futur.

En électronique et en informatique, il existe un grand nombre d'abréviations et de termes anglais, ils sont repris en tant que tels en fin d'article.

1. Développements des cartes à puces et leurs applications

1.1 Évolution des cartes à puces

Dès 1967, l'utilisation d'un composant électronique doté d'une mémoire dans une carte de crédit a fait l'objet de réflexions aux États-Unis, au Japon et en Europe, comme en témoignent **les très nombreux brevets** qui ont pavé le chemin de la carte à puces. Parmi les pionniers, on peut citer les Américains Pomeroy (1967), Ellingboe (1970), Castrucci (1971), Halpern (1972), le Japonais Arimura (1970), et les Français Moreno (1974), Ugon (1977) et Guillou (1979). La plupart de ces brevets n'ont pas donné lieu immédiatement à des réalisations, car ils anticipaient souvent sur les techniques disponibles.

En France, CII-Honeywell-Bull consacra des moyens de recherche importants, dès 1975, afin de définir l'architecture des composants et de trouver les moyens de réalisation des cartes. Ces recherches débouchèrent le **21 mars 1979** sur la **carte à microprocesseur**, après une coopération étroite avec Motorola. Ce fut la première carte à puces fonctionnant réellement. Appelée CP8, cette carte

était composée de deux puces, et elle fut essentielle pour prouver la faisabilité des concepts, convaincre les utilisateurs potentiels et lancer des expérimentations.

La Direction générale des télécommunications (DGT) commença alors à jouer un rôle moteur, multipliant les expériences, lançant dès 1980 des actions de normalisation et mettant à contribution ses propres organismes de recherche sur cette nouvelle technologie, dont le Centre national d'études des télécommunications (CNET), le Centre commun d'études de télédiffusion et télécommunications (CCETT) puis, plus tard, le Service d'études communes de la poste et de France télécom (SEPT). Sous cette impulsion, les sociétés Schlumberger et Philips se lancèrent à leur tour dans la course en explorant des voies différentes.

En 1980, le CCETT et Bull mirent au point la première carte d'abonnement « Antiope », et la première expérience mondiale de télépaiement à domicile fut réalisée en 1981 à Vélizy avec la Poste en utilisant la carte CP8 bi-puces.

Dans le monde de la téléphonie, les premières télécartes mises au point par Schlumberger et Thomson virent le jour dans les publiphones en 1983. En 1988, cinq ingénieurs de la société Thomson quittèrent la société pour fonder Gemplus Card International, qui obtiendra dès l'année suivante sa première commande d'un million de **télécartes de France Télécom**. Dopé par la formidable demande des opérateurs, notamment en France, au Mexique et en Chine, le marché atteint vite les 30 millions d'unités en 1990 pour dépasser 200 millions de cartes en circulation en 1992, et culminer à plus du milliard d'unités à la fin des années 1990. En 2018, plus de 10 milliards d'unités, tous types d'applications et facteurs de forme confondus, ont été mis sur le marché.

Au niveau technologique, la coopération entre Bull et Motorola se concrétisa par une deuxième étape clé en 1981, avec la naissance du **SPOM (*Self Programmable One Chip Microcomputer*)**, premier microcalculateur autoprogrammable monolithique pour carte à puces. C'est probablement l'événement déclenchant pour le marché de la carte à puces.

L'ensemble des banques françaises, représentées au sein d'un GIE (Groupement d'intérêt économique) Carte à mémoire, décidèrent en effet de tester dès 1982 les trois techniques développées par chacun des trois constructeurs : carte CP8 à microcalculateur monolithique de Bull à Blois, carte bi-puces de TRT-Philips à Caen et carte à logique câblée de Flonic-Schlumberger à Lyon. En 1985, le GIE devient GIE Carte bancaire et, à la suite de ces expériences, commanda finalement 16 millions de cartes à microcalculateur de type CP8. Ainsi, **la carte à puces bancaire** se généralisa en France en 1992. C'est incontestablement à partir de cette période que date le véritable démarrage du marché de la carte à puces.

À côté du développement du marché de la carte bancaire, un autre événement allait en effet se révéler particulièrement important par la suite : la démonstration de la **première carte SIM** en 1989 par Gemplus.

Dès 1987, treize pays européens s'étaient accordés sur les options de la future norme de téléphonie mobile européenne, suivant en cela les recommandations et spécifications du Groupe Spécial Mobiles, ou GSM, créé en 1982 par le CEPT. Initialement, fortement influencé par l'amélioration des systèmes analogiques, celui-ci avait finalement reçu, en 1986, la mission de spécifier un système numérique, qui devait être « aussi performant qu'un système analogique ». France Télécom était à l'origine de cette forte impulsion, et fut rapidement rejoint par ses homologues en Allemagne, Grande-Bretagne et Italie. Le système finalement proposé, dénommé GSM, adopta en matière de transmission radio le principe d'un accès multiple à répartition dans le temps, système dit « TDMA ». Il retint également l'idée des « sauts de fréquence » : l'émetteur et le récepteur changent de fréquence à intervalles définis au début de la communication. Finalement, il proposa un système d'identification/authentification des abonnés basé sur une carte à puce : le concept de carte SIM était né.

Le Conseil des Communautés européennes adopta en 1987 une recommandation (87/371/CEE), pour l'introduction coordonnée de communications mobiles terrestres publiques numériques paneuropéennes, et une directive (87/372/CEE), relative aux bandes de fréquence à réserver au système cellulaire. Les recommandations de l'époque spécifièrent, en outre, que les services commerciaux devaient démarrer en 1991. Le programme prit en fait un peu de retard. Ce n'est en effet qu'au 1er juillet 1991 qu'eut lieu en Belgique la première communication entre un mobile GSM et le réseau téléphonique fixe. Le démarrage s'accéléra rapidement en France, en Italie et dans les Pays Scandinaves. Dès 1992, tout en conservant son abréviation, le GSM fut rebaptisé « *Global System for Mobile Communications* ». Un changement de nom symbolique, illustrant parfaitement le passage du concept de laboratoire à celui de produit commercial. Dès 1995, Gemplus avançait des ventes cumulées de plus de 120 millions de cartes SIM…

Avec ces deux segments de marché bien établis, les **innovations technologiques** ont depuis connu une accélération sans précédent : première carte à coprocesseur RSA développée par Philips en 1992, introduction de la carte de santé SESAM-Vitale en France par Bull CP8 dès fin 1996, introduction des premiers systèmes de porte-monnaie électronique par Bull-CP8 en 1996, démonstration de la première carte Java par Schlumberger en 1997 (ce qui allait avoir une importance considérable pour l'ensemble du marché, nous y reviendrons plus loin), première carte Internet par Bull-CP8 en 2000, première carte à l'interface USB par Schlumberger en 2002, pour ne mentionner que quelques exemples significatifs.

Au-delà d'un signe indiscutable de grande vitalité de l'industrie, ces innovations annoncent aussi un changement radical de perspective : **la carte à puces est désormais de plus en plus intégrée** à l'environnement informatique personnel et professionnel de tout un chacun, constituant la ramification la plus fine du vaste réseau qui est en train de se constituer autour de la convergence des télécommunications, du grand public et de l'informatique. Elle constitue, et constituera de plus en plus, le lien privilégié entre un utilisateur et ses prestataires de services : clé d'accès aux différents types de réseau, mais aussi coffre-fort digital garantissant la sécurité informatique et les données privées des individus.

1.2 Applications et marchés de la carte à puces

Parmi les cartes à microcircuits, on peut distinguer deux familles de cartes :

– **les cartes à logique câblée**, dans lesquelles quelques fonctions simples sont fixées par les circuits électroniques interposés entre la mémoire non volatile et l'interface extérieure. Dans le bas de gamme, il existe aujourd'hui de multiples cartes à logique câblée centrées sur **le prépaiement de services** tels que le téléphone, le parking, le cinéma ou le lavage des voitures. Certaines cartes utilisent des composants standards contenant une simple mémoire non volatile à accès sérialisé et sans aucune protection sécuritaire. Ces cartes supportent essentiellement des fonctions d'identification utilisées principalement par des applications implémentant des programmes de fidélisation, plus rarement par des services à haute valeur ajoutée (ce fut par exemple longtemps le cas en Allemagne dans le domaine de la santé, jusqu'à ce que les niveaux de fraude constatés remettent en cause cette solution peu sécuritaire) ;

– **les cartes à microcalculateur**, quant à elles, possèdent la même structure qu'un ordinateur. Elles permettent non seulement de stocker des données mais aussi, et surtout, de traiter des informations de manière sécurisée : en effet, ces deux fonctions sont réalisées à l'aide d'un programme exécuté par un processeur central implanté sur un composant silicium. L'avantage évident de la carte à microcalculateur comparée à d'autres appareils électroniques – smartphone, ordinateur portable…, est que l'ensemble du composant est exclusivement dédié à la cryptographie et à la sécurité. Le logiciel embarqué est de taille réduite, ce qui permet une validation poussée et donc un niveau de fiabilité élevé. Les interfaces

externes sont réduites, ce qui facilite grandement le contrôle des entrées/sorties. Enfin, le chargement éventuel de code additionnel se fait via des protocoles standardisés à la sécurité éprouvée. Pour résumer, la simplicité et la spécialisation de la carte à microcalculateur comparée à d'autres appareils électroniques rendent possible la garantie d'un haut niveau de sécurité.

Les domaines privilégiés de la carte à puces sont tous ceux où l'utilisateur a besoin de posséder à proximité immédiate une **très grande sécurité de traitement des informations** : le microcalculateur et son programme embarqué permettent, par exemple, d'authentifier la carte et son porteur, de chiffrer et de déchiffrer des messages, ou de calculer des signatures électroniques apportant la preuve de l'effective réalisation d'une transaction licite. Les deux applications majeures de la carte à puces concernent d'une part l'authentification et l'identification sur les réseaux mobiles (GSM ou UMTS), et d'autre part le paiement électronique. La principale raison en est le compromis optimal offert par la technologie entre le coût de déploiement et le niveau de sécurité atteint pour une lutte efficace contre la fraude. Ces deux segments de marché (téléphonie mobile et paiement) représentaient en 2018 environ 85 % du marché global (en volume) de la carte à puces.

D'autres applications, actuellement en émergence, sont appelées à devenir à court ou moyen terme des relais de croissance pour le marché de la carte à puces : les applications gouvernementales (cartes d'identité, passeports électroniques, cartes de santé…), l'Internet des Objets incluant le M2M (Machine to Machine), la protection des smartphones et des ordinateurs d'entreprise, la télévision cryptée, la protection des droits d'auteur, les transports en commun… Signalons que, dans le cas de ces autres applications, la puce électronique est le plus souvent pourvue d'une interface sans contact, lui permettant, *via* une antenne, un transpondeur inclus dans la carte, la couverture ou la page de données du passeport, de communiquer par radio avec le monde extérieur.

2. Semi-conducteurs pour cartes à puces

Le cœur d'une carte à puces est constitué d'un composant électronique monolithique en silicium introduit dans l'épaisseur d'une carte en plastique. Avant d'aborder les deux grandes familles de composants utilisés pour les cartes à logique câblée et celles à microprocesseur, donnons un aperçu des technologies qui permettent de réaliser ces puces.

2.1 Technologies

À la genèse des cartes à puces, deux filières technologiques étaient en présence selon le type de transistor utilisé pour réaliser les circuits logiques. D'un côté, la **technologie dite bipolaire** réalisant un effet d'amplification de courant par la diffusion de porteurs majoritaires à travers les jonctions adjacentes de trois semi-conducteurs dopés. De l'autre, la **technologie MOS** (*Metal Oxide Semiconductor*) fondée sur des transistors unipolaires utilisant la conduction d'un seul type de porteurs dans un mince canal contrôlé par une électrode isolée. Suivant en cela la logique économique reprise par l'ensemble de l'industrie électronique, l'industrie de la carte à puces s'est ralliée massivement, dès l'origine, à la filière MOS. Les raisons principales sont, d'une part, des puissances consommées beaucoup plus faibles qu'en bipolaire, et d'autre part de très grandes capacités d'intégration. Au cours des deux dernières décennies, un autre avantage déterminant est l'évolution de la **technologie CMOS** (*Complementary MOS*), qui se traduit par une très faible consommation et une bonne immunité au bruit. La figure **1** montre le schéma de principe d'un transistor CMOS.

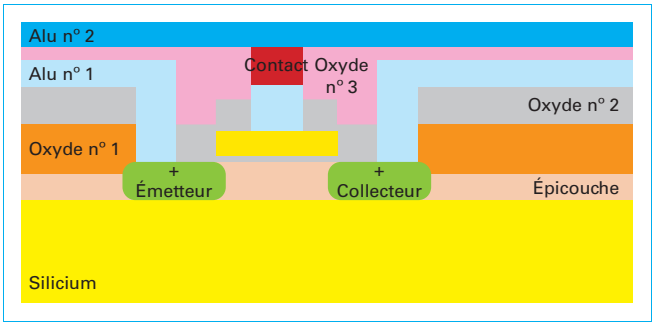


Figure 1 – Schéma de principe d'un transistor CMOS

Dans les cartes à puces, il faut pouvoir mémoriser des informations confidentielles, y compris en l'absence d'alimentation des circuits. Cette caractéristique constitue l'un des éléments fondamentaux des composants pour cartes à puces : en plus d'un microprocesseur spécialisé, les cartes doivent embarquer **différents types de mémoires spécialisées**. Celles-ci sont utilisées selon la nature de la mémorisation considérée, la vitesse de fonctionnement et la volatilité des informations stockées.

2.1.1 Mémoires à accès aléatoire

Les **mémoires à accès aléatoire**, appelées RAM (*Random Access Memory*), sont utilisées en tant que registres de travail temporaire et perdent leurs informations dès qu'elles ne sont plus alimentées. Les mémoires RAM dynamiques (DRAM) sont bâties à partir d'un seul transistor et il faut renouveler leur contenu périodiquement, tandis que les mémoires RAM statiques (SRAM), à quatre ou six transistors, possèdent deux états stables permettant de stocker un élément binaire. Dans les cartes à puces, les registres de travail sont des mémoires SRAM, avec un impact important sur le prix des composants, puisqu'une telle cellule occupe environ 20 fois plus de place qu'une cellule ROM (*Read Only Memory*). L'optimisation des coûts explique donc les tailles RAM relativement faibles que l'on trouve dans ces composants : les tailles typiques de mémoires RAM embarquées varient entre quelques kilooctets (ko) et quelques dizaines de kilooctets.

2.1.2 Mémoires non volatiles

Les **mémoires non volatiles** (NVM) gardent les informations en l'absence d'alimentation électrique. Il existe deux types principaux de mémoires NVM disponibles sur les composants pour carte à puce :

– **les mémoires mortes** (ROM), inaltérables, contiennent des informations permanentes telles que les programmes. Elles ne sont accessibles qu'en lecture. L'inscription de la ROM est réalisée par masquage ou par implantation ionique dans le silicium pendant la fabrication du circuit intégré ; les tailles des mémoires ROM embarquées dans le composant pour carte à puces varient en général entre quelques dizaines et quelques centaines de kilooctets ;

– **les mémoires mortes programmables** (PROM) peuvent être programmées (ou écrites) par l'utilisateur. Dans la technologie MOS, les niveaux d'isolement sont tels que l'on peut enregistrer des informations en piégeant des charges électriques dans une électrode flottante. L'évacuation de ces charges permet d'effacer ce type de mémoire. À l'origine, l'effacement était réalisé à l'aide d'un rayonnement ionisant EPROM (*Erasable Programmable ROM*) ne permettant qu'un petit nombre de re-programmations des cellules mémoire. Aujourd'hui, la quasi-totalité des composants utilisent des mémoires à effacement et re-programmation par application d'une tension électrique, ou EEPROM (*Electrically Erasable Programmable ROM*). Pour programmer une cellule, on fait circuler un courant intense entre la source et le drain. Certains électrons acquièrent une énergie leur permettant d'atteindre la grille

flottante, ils y sont alors piégés. Lorsque la charge piégée est suffisante, elle masque le champ électrique induit par la grille et le transistor est bloqué. Le courant de fuite étant très faible, cette charge peut se conserver très longtemps. Il est également possible de décharger la grille flottante par effet tunnel en appliquant des tensions suffisamment élevées entre la grille, la source et le drain. Cet effacement est plus rapide et n'impose pas de retirer le circuit du système dans lequel il est installé. La différence entre les EEPROM et les mémoires vives réside essentiellement dans la vitesse d'écriture. Le cycle d'écriture d'une EEPROM est environ 1 000 fois plus long que celui d'une RAM. Les mémoires EEPROM sont par ailleurs trois fois plus encombrantes que les mémoires SRAM. Les capacités des mémoires EEPROM embarquées dans les cartes à puces se situent entre 2 et 400 ko, avec, pour des raisons essentiellement technologiques, une limite supérieure prévisible autour de 512 ko à 1 Mo ;

Les limitations des mémoires EEPROM ont favorisé leur remplacement par **des mémoires flash**. Celles-ci fonctionnent par stockage d'électrons dans une couche mince de polysilicium en suspension dans un oxyde, sous une grille de contrôle « *on-off* » d'un transistor. Le principe de lecture de la cellule flash est simple : il s'agit de mesurer si une tension appliquée à la grille de contrôle allume ou non le transistor. L'écriture est en revanche plus complexe et s'effectue en deux phases : tout d'abord, il faut enlever les charges d'un bloc de cellules mémoires (un bloc peut comporter plusieurs milliers de transistors), puis la cellule est programmée par injection (ou non) d'électrons dans la grille flottante. Cette opération nécessite une énergie élevée pour faire passer les électrons à travers la barrière d'oxyde isolante. Elle endommage à la longue la couche isolante, ce qui explique l'altération des performances des mémoires flash avec le temps. Les mémoires flash présentent l'avantage d'un faible encombrement (la densité est comparable à celle de la mémoire ROM) et des performances en écriture sensiblement meilleures que celles des mémoires EEPROM. Leur granularité d'accès plus faible que celle des EEPROM pose toutefois des problèmes de programmation complexes. L'introduction des mémoires flash dans les composants cartes à puces est cependant devenue une réalité depuis 2005, et c'est imposé depuis comme un véritable standard.

L'industrie des semi-conducteurs et les industriels de la carte à puces ont mis en place des concepts de sécurité spécifiques sur les mémoires flash pour les amener au même niveau que les mémoires ROM et obtenir les mêmes certifications de sécurités par des laboratoires indépendants. Il en résulte qu'il est même possible de remplacer les mémoires ROM complètement.

Cela ouvre la possibilité non seulement de corriger une erreur dans le système d'exploitation (OS) de la puce (*code patching*), mais de reporter l'implémentation du système d'exploitation et de son verrouillage à l'étape de la personnalisation. Les mémoires flash apportent donc flexibilité et rapidité pour la mise en place de nouvelles applications sur le terrain.

Le tableau **1** résume les principales propriétés des mémoires utilisables dans les composants pour cartes à puces.

Le lecteur pourra aussi consulter l'article [E 2 430].

2.2 Composants en logique câblée

Ces composants, généralement très simples, sont des NVM à accès sériel synchrone, contrôlés par des circuits logiques interconnectés entre la mémoire et l'interface externe. Les commandes disponibles sur ces composants sont très limitées et figées par construction. Les cartes à logique câblée sont principalement utilisées comme jetons électroniques ou comme identifiants. L'exemple le plus répandu était la **télécarte** utilisée dans les téléphones publics et dont le composant, fabriqué à plusieurs centaines de millions d'exemplaires, détenait longtemps le record du monde en termes de quantité, maintenant supplanté par les cartes SIM.

Dans ce domaine, il n'y a malheureusement aucun véritable standard et les composants offerts sont très souvent incompatibles entre eux.

La famille des jetons électroniques est déjà assez riche et s'est développée en deux générations.

■ La **première génération** (1G) correspond au lancement de la télécarte en France et en Allemagne, elle se compose essentiellement de mémoires sérialisées possédant une zone protégée par un fusible ;

Dans la première télécarte, la mémoire EPROM possède 256 bits, avec une zone protégée de 96 bits qui stocke une référence. Par contre, la carte allemande, développée plus tard, contenait 416 bits de mémoire EEPROM recyclable 64 fois, incluait 208 bits de mémoire de travail et un contrôle d'accès par un code. Ces deux réalisations ont engendré deux lignées d'interfaces électriques incompatibles entre elles : d'une part, une interface nécessitant 8 contacts externes, utilisée par France Télécom, et d'autre part, deux interfaces à 8 et 6 contacts utilisées par la Bundespost (DBP). La figure **2** illustre les trois types d'interfaces utilisées par France Télécom, Bundespost et recommandées par la norme ISO, ainsi que leurs commandes associées.

La méthode d'accès à ces mémoires consiste à adresser bit à bit chaque cellule comme dans un registre à décalage. La remise à zéro du circuit permet de se positionner sur la première adresse mémoire et la lecture se fait sur le plot OUT, après la remontée du signal d'horloge.

Lorsque le signal RST de remise à zéro n'est plus actif, une impulsion d'horloge sur CLK incrémente le compteur d'adresse du composant. Les commandes de lecture, d'écriture ou d'effacement correspondent aux combinaisons de signaux des figures ***2d*** et ***2e***, et sont prises en compte sur un front montant de l'horloge. Dans le cas de l'interface de type FT, il faut utiliser une combinaison de deux signaux, RST et B, tandis que dans la première télécarte allemande (DBP), il s'agit des signaux RST et PROG. Dans le cas de l'interface à 5 contacts, inspirée par le standard ISO, toutes les commandes sont assurées par le plot RST. Dans ce cas, une impulsion sur RST en maintenant CLK à zéro permet de passer en mode programmation, tout en restant sur l'adresse sélectionnée, et l'impulsion suivante sur CLK permet de programmer le bit correspondant.

■ Dans les composants de **seconde génération** (2G) apparaissent deux caractéristiques nouvelles. D'une part, une fonction

Tableau 1 – Comparaison des différents types de mémoires pour cartes à puces

	EEPROM	Flash	MRAM	PCRAM
Taille relative versus DRAM	5-10	0,25-1	1-3	0,8-2
Granularité	Octet	Bloc/secteur	Bit	Bit
Endurance (en cycles)	10^6	10^5	$> 10^{14}$	10^{12}
Temps d'écriture (Programmation/effacement)	$ms.ms^{-1}$	$\mu s.ms^{-1}$	< 100 ns	< 100 ns
Puissance en écriture	10 V x 100 µA	5 V x 1 mA	1,8 V x 10 mA	3 V x 1mA

d'authentification dynamique, qui délivre un résultat sur 4 bits lorsque l'on fournit à la carte une donnée aléatoire de 64 bits et, d'autre part, une mémoire de travail constituée par des compteurs fonctionnant suivant un mécanisme de boulier. La figure **3** donne l'organisation de la mémoire des composants ST 1335 et 1336 de STMicroelectronics, qui rassemblent pratiquement toutes les caractéristiques rencontrées dans les divers autres composants à logique câblée.

Une première zone protégée de 64 bits contient un code d'identification inscrit en usine et des informations sur l'application introduite par l'émetteur. Cinq registres de 8 bits constituent les compteurs et sont associés à une autre zone de quatre registres d'indicateurs témoins qui permettent les reprises, en cas de rupture de séquence à la suite d'une extraction intempestive. La carte peut aussi contenir une signature électronique attestant les droits de la carte et une clé cryptographique pour l'authentification.

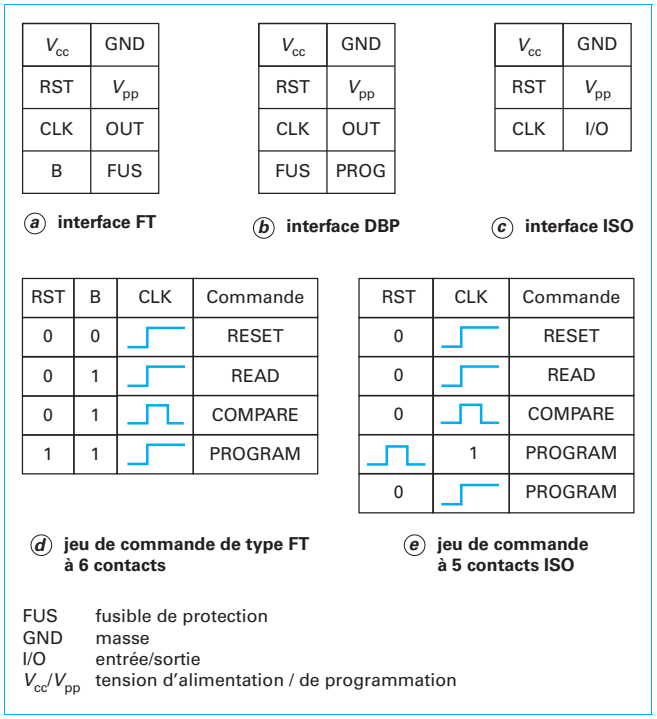


Figure 2 – Interfaces et commandes associées

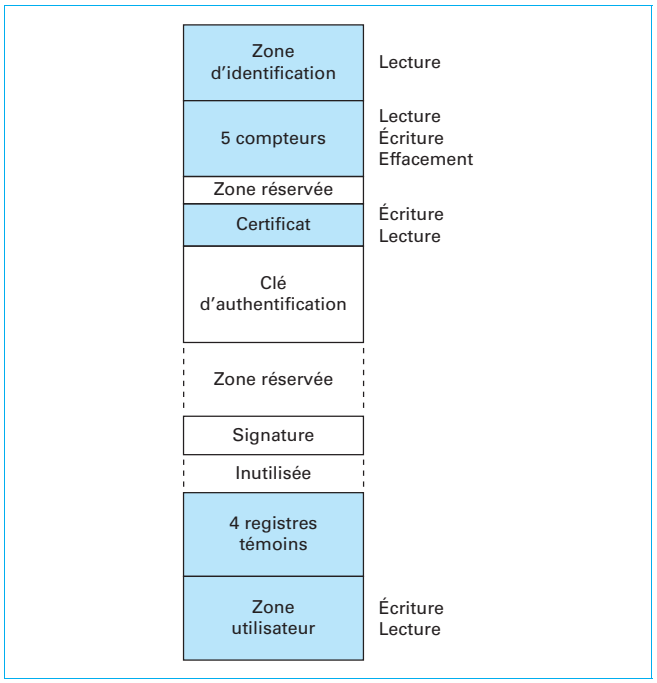


Figure 3 – Organisation mémoire (34 x 8 bits) (composants ST 1335 et 1336 de STMicroelectronics)

2.3 Microcalculateurs

Les microcalculateurs pour cartes à puces sont **de véritables ordinateurs intégrés sur un seul substrat de silicium** : ce sont donc essentiellement des machines à programme enregistré qui sont architecturées autour d'un (parfois deux) bus de données. Ils contiennent les différents types de mémoire, et les organes d'entrée-sortie, qui assurent les dialogues avec le monde extérieur. Le programme de fonctionnement est généralement contenu dans la ROM, EEPROM ou FLASH, tandis que la RAM contient les registres de travail nécessaires aux divers traitements internes.

Les mémoires non volatiles intégrées sur le composant permettent de reprogrammer la mémoire, à des fins de mise au point ou d'évolution, voire pour les architectures les plus modernes de rajouter des programmes après la délivrance de la carte (par exemple, dans le cas des architectures Javacard). Les microcalculateurs pour cartes à puces intègrent encore de nos jours le concept de SPOM (*Self Programmable One Chip Microcomputer*), architecture essentiellement sécuritaire, qui permet au microcalculateur de modifier lui-même un programme contenu dans sa propre NVM sans intervention du monde extérieur (figure **4**). À cette caractéristique viennent s'ajouter divers dispositifs de sécurité qui empêchent les

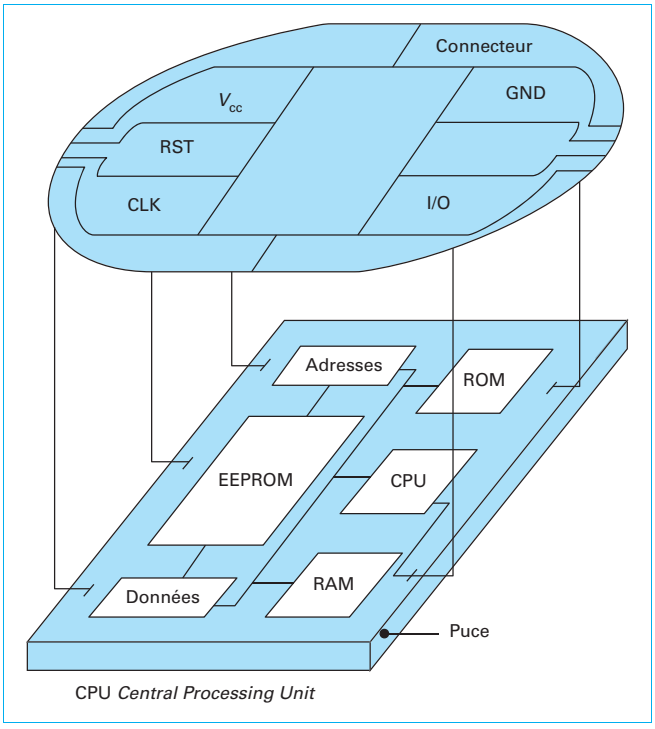


Figure 4 – Microcalculateur SPOM et ses interconnexions

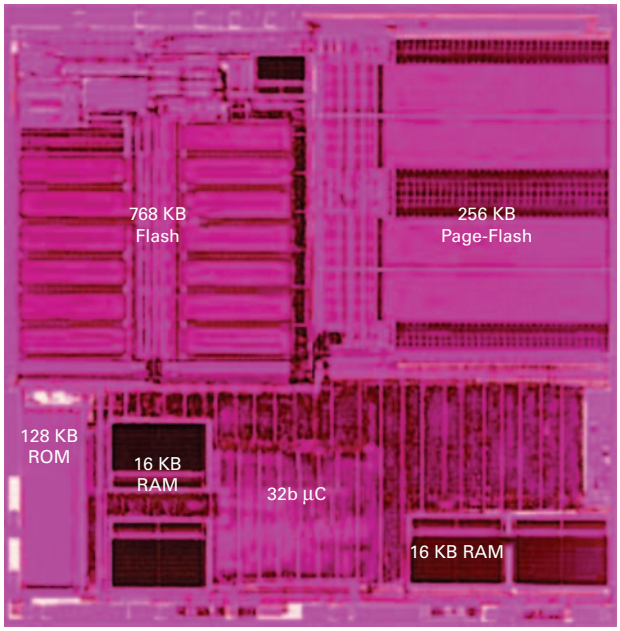


Figure 5 – SPOM ST 22FJ1M de STMicroelectronics

intrusions dans le composant, telles que celles qui sont décrites au paragraphe 3. La figure **5** montre un exemple de composant : le SPOM ST 22FJ1M de STMicroelectronics réalisé en technologie CMOS 0,18 µm, qui intègre des mémoires flash, en remplacement des traditionnelles mémoires EEPROM.

Comme les microcalculateurs sont des machines programmables, ces composants sont **adaptables à l'infini**, pour répondre à une application donnée, ou évoluer par une simple modification du programme enregistré dans la ROM, ou même dans la NVM, dès lors qu'un certain nombre de précautions sécuritaires sont prises. Bien entendu, les données applicatives sont généralement mémorisées dans la NVM. En outre, la capacité de traitement des microprocesseurs permet de rendre l'interface externe indépendante des applications, et donc normalisable. L'interface électrique, les protocoles d'échanges et le jeu de commandes de ces cartes font l'objet de la norme internationale ISO IEC 7816, dont la description est faite en annexe dans la partie documentation de l'article.

3. Cryptologie et sécurité

La cryptologie peut se définir comme la science de la protection logique de l'information. Elle constitue, avec la sécurité physique des composants et la sécurité des logiciels, une dimension essentielle de la sécurité.

La cryptologie englobe la **cryptographie**, qui est l'art d'authentifier, de signer, de vérifier, de chiffrer, de déchiffrer… l'information, et la **cryptanalyse**, qui est l'art de « casser » les solutions proposées par la cryptographie.

La cryptographie repose sur **des briques de base qui sont des algorithmes** *ad hoc*, dont la conception est basée sur des principes mathématiques éprouvés. La plupart des algorithmes cryptographiques utilisés sont des algorithmes standardisés qui ont fait l'objet pendant plusieurs mois d'une cryptanalyse poussée par la communauté internationale des experts en cryptologie.

Il est rare qu'un algorithme seul suffise à bâtir une solution sûre et utilisable dans le monde réel. Dans la très grande majorité des cas d'usage, une solution requiert la collaboration de plusieurs entités, qui doivent alors échanger l'information en respectant des propriétés de sécurité : confidentialité, authenticité, anti-rejeu, non-répudiation… Les entités mettent alors en œuvre des algorithmes cryptographiques chacune de leur côté pour échanger l'information suivant des protocoles élaborés afin de fournir les propriétés recherchées. On parle alors de protocoles cryptographiques, dont la spécification doit être définie avec précaution. Comme pour les algorithmes, une revue des spécifications par les experts du domaine est indispensable pour éviter l'introduction de faiblesses dans les protocoles. En plus des briques de base que sont les algorithmes cryptographiques, ces protocoles reposent également sur l'utilisation de nombres aléatoires, dont la qualité constitue un élément essentiel de leur sécurité. L'entropie et l'imprédictibilité sont les deux propriétés à considérer pour évaluer la qualité des nombres aléatoires. Certaines agences de sécurité nationales comme le NIST et le BSI ont défini des conceptions standards de source d'aléas.

Si la sécurité mathématique des algorithmes et des protocoles cryptographiques est un prérequis indispensable à la construction de solutions sûres, un autre volet de la sécurité est **la robustesse de mise en œuvre des logiciels et des matériels utilisés**. En effet, une mise en œuvre naïve, logicielle ou matérielle, d'un algorithme standardisé permettra par exemple à un attaquant de retrouver la clé cryptographique utilisée, et donc au final de compromettre l'information échangée. Au-delà de la mise en œuvre de l'algorithme cryptographique lui-même, le logiciel utilisant le service cryptographique doit également se protéger de différents types d'attaques. Dans le cas contraire, un logiciel malveillant pourrait par exemple utiliser le service cryptographique à la place du logiciel légitime, ou bien en extraire la clé.

Une solution sûre repose donc sur des entités qui échangent de l'information suivant **un protocole cryptographique éprouvé**, mis en œuvre par les différentes entités grâce à des logiciels et des matériels convenablement protégés qui exécutent des algorithmes cryptographiques et des générations de nombres aléatoires conformes à l'état de l'art.

Enfin, pour qu'une solution soit opérationnelle, les différentes entités doivent manipuler des clés cryptographiques dont la génération et la gestion doivent également faire l'objet d'une grande attention. La gestion des clés – génération, distribution, attestation, révocation… – est ainsi un élément déterminant de la sécurité d'une solution.

Dans ce panorama d'exigences, la carte à puces propose des caractéristiques idéales de sécurité et de coût pour participer à la mise en œuvre de solutions sûres.

3.1 Cryptographie

3.1.1 Principes de la cryptographie

La cryptographie repose sur l'utilisation de briques de base – les algorithmes cryptographiques – validées au préalable par la communauté internationale des experts en cryptologie. Aujourd'hui, les organismes de standardisation exigent qu'un algorithme soit accompagné de sa preuve mathématique de sécurité pour prétendre à rentrer dans un standard. Cette preuve est alors étudiée pendant plusieurs mois par les experts à travers une cryptanalyse de l'algorithme. L'algorithme est standardisé si et seulement si les conclusions de l'étude valide la preuve apportée.

Ces briques de base, par exemple les algorithmes AES, RSA, EC-DSA, SHA-256…, sont utilisées quotidiennement par chacun d'entre nous lors d'une connexion internet sécurisée, d'une transaction bancaire, d'une vérification de passeport électronique, etc.

Une liste exhaustive des algorithmes utilisés n'est pas l'objet de ce document, mais le lecteur peut se référer aux liens ci-dessous s'il souhaite obtenir un panorama des algorithmes standardisés.

Les algorithmes sont suivis avec attention par les agences nationales de sécurité qui mesurent leur érosion. Deux facteurs contribuent à l'érosion d'un algorithme : l'augmentation de la puissance de calcul disponible pour les « attaquants » et les progrès de la cryptanalyse. Les agences de sécurité (NIST (États-Unis), ANSSI (France), BSI (Allemagne), ENISA (Europe)…) ou les réseaux d'experts (ECRYPT (Europe)) publient régulièrement leur position actualisée sur la solidité et l'utilisation des algorithmes voici trois liens qui en attestent :

– https://www.iacr.org/newsletter/v21n2/ecrypt.html
– https://www.ssi.gouv.fr/uploads/2014/11/RGS_v-2-0_B1.pdf
– https://www.keylength.com/

La force d'un algorithme, c'est-à-dire sa capacité à résister à toute cryptanalyse connue, est mesurée en nombre de bits. Cette mesure permet une comparaison aisée de la force relative des différents algorithmes. Elle évolue au cours du temps pour un algorithme en fonction de son érosion. La force d'un algorithme dépend bien évidemment de la taille de la clé utilisée. Aussi, la force d'un algorithme est toujours associée à l'algorithme considéré et à une taille de clé.

Il existe différents algorithmes qui sont utilisés en fonction du service recherché : chiffrement/déchiffrement, authentification d'un message, chiffrement authentifié, chiffrement à clé publique, signature, échange de clés, dérivation de clés…

Le chiffrement des messages consiste par exemple à transformer un message de telle sorte que seule une entité autorisée puisse le retrouver. La fonction de transformation constitue l'algorithme cryptographique, dont le secret réside dans des paramètres appelés **clés**. Lorsque l'on déchiffre le message, on réalise l'opération inverse en utilisant ces clés. Dans les cartes à puces, la cryptographie met en œuvre divers mécanismes qui ont pour but d'assurer soit la confidentialité des informations, soit l'authentification des cartes ou des utilisateurs, soit encore la signature des messages. L'ensemble des moyens mettant en œuvre la cryptographie forme un **cryptosystème**.

3.1.2 Cryptosystèmes symétriques

Le schéma simplifié d'un cryptosystème est représenté figure **6**. À l'émission, le message est transformé par l'algorithme A qui est fonction de la clé de chiffrement K_e . À la réception, les informations reçues sont déchiffrées à l'aide de l'algorithme inverse A^{-1} utilisant une clé K_r. Notons que le déchiffrement implique **un algorithme réversible**, ce qui n'est pas le cas de l'authentification. En effet, certains mécanismes d'authentification utilisent un algorithme dans le même sens aux deux extrémités avec la même clé.

Le cryptosystème est dit **symétrique** lorsque $K_e = K_r$. Les algorithmes utilisés dans ces cryptosystèmes symétriques sont très nombreux. Ils peuvent être privés ou publics, mais le choix d'un algorithme pour carte à puces dépend essentiellement du service de sécurité attendu, de la performance, et surtout du coût des ressources nécessaires à son implantation (tailles de RAM et de ROM). En effet, l'utilisation d'algorithmes encombrants augmente très vite le prix des cartes. Il est également important de tenir compte de la réglementation concernant l'usage de la cryptographie qui est variable d'un pays à l'autre.

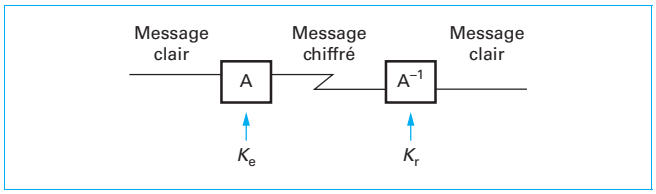


Figure 6 – Schéma simplifié d'un cryptosystème

Le plus ancien des algorithmes symétriques utilisé massivement est certainement le **DES (*Data Encryption Standard*)** normalisé par le NIST (*National Institute of Standard and Technology*) aux États-Unis. Développé dans les années 1970, le DES n'a pas été conçu pour une implémentation programmée et ne correspond probablement pas au meilleur compromis souhaitable entre performances et ressources utilisées. Cependant, sa relative simplicité d'implémentation, ainsi que sa sécurité intrinsèque l'ont rendu pratiquement incontournable dans les cartes à microcalculateurs. Utilisant une longueur de clés de 56 bits, et prenant en entrée des blocs de 64 bits, le DES tend à devenir obsolète et à être remplacé par un nouvel algorithme, connu sous le nom d'AES ou « *Advanced Encryption Standard* », issu d'un appel à candidatures international lancé en janvier 1997 par le NIST.

L'algorithme AES, adopté par le NIST en 2001, prend en entrée un bloc de 128 bits (16 octets), pour des longueurs de clés de 128, 192 ou 256 bits. Les 16 octets en entrée sont permutés, selon une table définie au préalable. Ces octets sont ensuite placés dans une matrice de 4 x 4 éléments, et ses lignes subissent une rotation vers la droite, variant en fonction du numéro de la ligne. Une transformation linéaire est ensuite appliquée sur la matrice : elle consiste en la multiplication binaire de chaque élément de la matrice par des polynômes issus d'une matrice auxiliaire, suivant les règles algébriques propres à un corps fini (en l'occurrence le corps fini à 128 éléments soit GF (2^8) pour *Gallois field*). Cette transformation linéaire garantit une meilleure propagation des bits dans la structure sur plusieurs tours. Finalement, un « ou-exclusif », ou XOR, entre la matrice et une autre matrice, permet d'obtenir une matrice intermédiaire. Ces différentes opérations sont répétées plusieurs fois et définissent un « tour » ou itération de l'algorithme. Pour une clé de 128, 192 ou 256, l'AES nécessite respectivement 10, 12 ou 14 tours. Aucune véritable faiblesse cryptographique n'a été décelée à ce jour dans l'algorithme AES, adopté comme système de chiffrement des données non sensibles du gouvernement américain. Peu gourmand en mémoire, l'AES se prête particulièrement bien à une implémentation dans une carte à puces.

Les algorithmes de chiffrement constituent en fait une boîte noire à trois canaux : deux d'entrée (la clé et l'information d'origine) et un de sortie, dont la largeur est généralement un multiple de 64 bits. On peut utiliser **deux principaux modes de fonctionnement** :

– le mode de base ECB (*Electronic Code Book*) qui transforme simplement les n x 64 bits d'entrée en un message de même longueur ;

– le mode CBC (*Cipher Bloc Chaining*) qui combine les blocs de données en les chaînant : le $n^{ième}$ cryptogramme, C_n, d'un message M est en fait calculé à partir du $n^{ième}$ bloc de message, M_n, et du $(n-1)^{ième}$ cryptogramme, C_{n-1}, par la relation :

$$C_n = EK\left(C_{n-1} \text{Å } M_n\right)$$

avec K clé d'encryption,
E algorithme utilisé,
Å représente le « ou » exclusif.

Cela permet de chiffrer différemment des blocs identiques. Ce mode est également utilisé pour générer la signature d'un message en conformité avec la norme ANSI X9.9.

Un exemple d'**authentification dynamique** est donné figure **7**. Le vérifieur choisit un nombre E au hasard, le chiffre avant de l'envoyer à la carte, et demande à cette dernière de déchiffrer le message M. Si la carte est authentique, elle est la seule à pouvoir retrouver le nombre E en utilisant sa clé secrète (ici l'algorithme est symétrique donc $K_1 = K_2$).

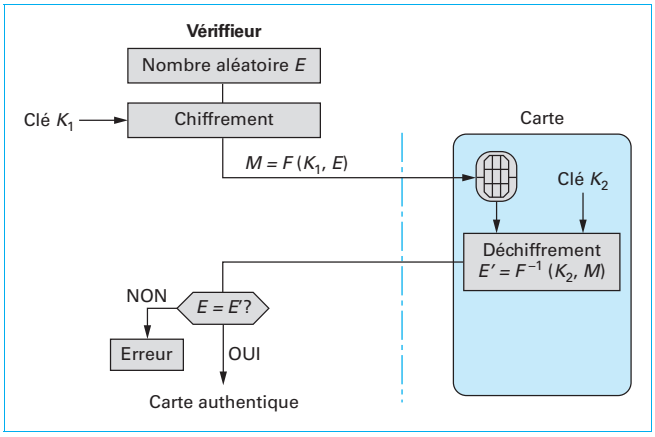


Figure 7 – Mécanisme d'authentification

On ne peut faire ici une présentation exhaustive des algorithmes symétriques, mais il faut attirer l'attention sur les problèmes posés de gestion de clés. En effet, lorsqu'un grand nombre d'utilisateurs font partie d'un réseau, il faut que chacun d'entre eux possède une clé personnalisée, car une clé unique constituerait une menace pour tout le système dans le cas où elle serait compromise. Comme il est peu pratique et risqué de stocker toutes les clés, la méthode consiste à les diversifier à partir d'une clé maître et d'un identifiant de chaque carte. Les clés maîtres doivent être particulièrement protégées, et elles sont parfois contenues dans un module de sécurité ou une carte « mère » possédée par l'émetteur des cartes.

3.1.3 Cryptosystèmes asymétriques

Lorsque K_e est différent de K_r, le cryptosystème est dit asymétrique. L'idée de base, qui date de 1976, suppose qu'il existe des algorithmes cryptographiques tels qu'il soit impossible, au sens de la complexité des calculs, de retrouver une des deux clés en connaissant l'autre.

Ce n'est qu'en 1978 qu'un tel algorithme a vu le jour : il s'agit du **RSA**, du nom de ses inventeurs Rivest, Shamir et Adleman. Le RSA est fondé sur la difficulté de factoriser des grands nombres. N étant un nombre composite, dont les facteurs premiers sont p et q, le chiffrement RSA du message M est donné par :

$$C = M^e \bmod N$$

On montre que la fonction de déchiffrement est :

$$M = C^d \bmod N$$

et que la relation entre les deux exposants e et d est :

$$ed = 1 \bmod (p-1)(q-1)$$

Si N est suffisamment grand, on peut donc le publier, ainsi que l'un des deux exposants e ou d, et garder p et q secrets. Dans ces conditions, quelqu'un ne connaissant pas p et q ne peut calculer e en fonction de d, et inversement. Cela permet de divulguer l'une des deux clés en la mettant, par exemple, dans un annuaire public. Les propriétés asymétriques de ces algorithmes sont très intéressantes pour chiffrer, authentifier ou signer des messages. Ainsi :

– pour assurer la confidentialité d'une information, toute personne connaissant N et e peut chiffrer un message. Seul le destinataire détenant la clé secrète pourra le déchiffrer ;

– n'importe qui connaissant la clé publique peut authentifier une carte grâce au schéma de la figure **7** : le vérifieur choisit un nombre E au hasard, le chiffre à l'aide de la clé publique K_1 liée à la carte et demande à cette dernière de déchiffrer le message. Si la carte est authentique, elle est seule à pouvoir retrouver le nombre E en utilisant sa clé secrète K_2. Si l'on veut utiliser cette méthode avec un temps de calcul raisonnable, il faut donc que le micro-calculateur de la carte puisse calculer rapidement les multiplications modulaires ;

– toute personne peut vérifier la signature d'un émetteur sans détenir son secret.

Bien que les cryptosystèmes à clé publique soient très séduisants, il faut aussi prendre un **certain nombre de précautions** qui peuvent compliquer l'apparente simplicité de la gestion des clés.

Que les clés publiques soient dans un annuaire ou qu'elles soient transmises, il est nécessaire de les signer à l'aide de la clé secrète de l'émetteur pour éviter que des intrus pénètrent le système en publiant leur propre clé. Par ailleurs, dans un réseau réel, il faut calculer autant de nombres composites que d'utilisateurs et garantir un moyen sûr de transmission des clés secrètes de l'autorité vers les abonnés et des clés publiques des abonnés vers l'autorité. Les cartes à puces elles-mêmes peuvent constituer un excellent vecteur de transmission, mais on voit que cette fonction de distribution conditionne la viabilité du système.

Enfin, nous avons signalé l'importance de la longueur de la clé qui dépend du niveau de sécurité que l'on veut atteindre, de la durée d'utilisation des clés et du risque encouru. Il y a donc lieu de tenir compte des progrès en matière de factorisation. Il est aujourd'hui admis que, pour la prochaine décennie, une application de sécurité moyenne basée sur le RSA ne devrait pas descendre au-dessous de 1 024 bits, et que, dès qu'il s'agit de haute sécurité, il faut passer à 2 048 bits. Si le RSA constitue aujourd'hui le système à clés publiques le plus employé dans les cartes à microcalculateurs pour les transactions bancaires, d'autres systèmes, tels les cryptosystèmes basés sur des courbes elliptiques sur des corps finis, sont utilisés dans les passeports électroniques et les documents d'identité.

3.1.4 Calcul quantique et cryptographie post-quantique

En 2016, IBM a mis au point un processeur 5-qubit permettant de faire les tous premiers essais de calcul quantique.

En informatique quantique, un **qubit**, parfois écrit qbit, est l'état quantique qui représente la plus petite unité de stockage d'information quantique. C'est l'analogue quantique du bit.

Un ordinateur ou un accélérateur quantique de plusieurs centaines de qubits serait capable de résoudre certains problèmes bien plus rapidement qu'un ordinateur traditionnel et permettrait alors de réduire de manière exponentielle les temps de calcul pour des applications médicales, d'intelligence artificiel et de big data. Mais, il pourrait aussi être utilisé efficacement pour la cryptanalyse de certains algorithmes cryptographiques.

Cela pourrait avoir des répercussions sur la cryptographie actuelle, en particulier sur celle à clé publique. Les algorithmes à base de courbes elliptiques ne seraient par exemple plus suffisamment sûrs. La **cryptographie à clés symétriques** (par exemple l'AES-256) semble aujourd'hui **plus résistante à une cryptanalyse** basée sur le calcul quantique que la cryptographie à clés publiques.

Afin d'éviter toute crise de sécurité dans le futur, il est d'ores et déjà nécessaire de travailler sur des algorithmes cryptographiques résistants aux calculs quantiques. De tels algorithmes sont dit **« post-quantiques »**. Le NIST a décidé en 2016 de lancer officiellement des travaux de recherche pour la définition de d'algorithmes post-quantiques, en particulier pour offrir une alternative aux algorithmes à clés publiques actuels. Les premiers produits cartes à puce incluant une sécurité post-quantique sont attendus pour 2020-2022.

D'ores et déjà, certaines agences nationales recommandent le déploiement de **solutions crypto-agiles**. La crypto-agilité consiste

à prévoir dans les solutions déployées la possibilité de changer les briques de bases sans remettre en cause l'ensemble de la solution. Ainsi, un algorithme mis à mal par le calcul quantique pourra être remplacé par son équivalent post-quantique.

Une autre utilisation des propriétés quantiques est la **distribution de clés**. Basé sur le principe qu'une mesure effectuée par un attaquant modifie nécessairement l'état quantique d'une particule, l'échange de clés quantiques offre des propriétés de sécurité basés pour partie sur ce principe de base de la mécanique quantique. L'échange de clés quantiques est déjà mis en œuvre en pratique dans plusieurs pays, aujourd'hui sur des distances relativement courtes.

3.2 Sécurité physique et logique des cartes à puces

Les microcalculateurs pour cartes à puces assurent la confidentialité et l'intégrité des informations sensibles, telles que les clés cryptographiques ou les mots de passe. Ils assurent également l'exécution non modifiable des programmes embarqués dans la puce.

3.2.1 Attaques

Une carte à puce est en fait conçue pour résister à tous types d'attaque, sauf bien entendu dans le cas d'une mise en œuvre de moyens très conséquents. Les attaques sont en général classifiées comme suit : logiques ou physiques.

3.2.1.1 Attaques logiques

Les attaques logiques consistent à compromettre une donnée ou une fonction en utilisant les interfaces physiques normales du dispositif. Ces attaques utilisent par exemple l'injection de code malicieux, le débordement de pile ou de tableau, l'exploitation de bogues ou de mauvaises configurations, l'absence de tests aux limites… Elles peuvent s'envisager en local, lorsque l'attaquant doit se trouver à proximité immédiate du dispositif, ou à distance, lorsque l'attaquant échange avec le dispositif à travers un réseau.

3.2.1.2 Attaques physiques

Les attaques physiques consistent à utiliser des moyens matériels pour compromettre la confidentialité ou l'intégrité des informations sensibles contenue dans le dispositif ou encore pour modifier l'exécution des programmes embarqués. Les attaques physiques nécessitent la présence du dispositif à proximité de l'attaquant et elles se classifient comme suit :

– attaques par canaux cachés,
– attaques par perturbation,
– attaques intrusives.

Les attaques par canaux cachés consistent à mesurer une grandeur physique lorsque le dispositif est en fonctionnement, puis à analyser les mesures effectuées pour en déduire des informations sensibles, par exemple une clé cryptographique. Les grandeurs physiques mesurées sont couramment le temps d'exécution d'un processus, la consommation de courant du dispositif, son rayonnement électromagnétique.

La figure **8** montre une mesure en courant effectuée lors d'un calcul RSA.

L'analyse des mesures peut recourir à des techniques statistiques élaborées, qui, le cas échéant, permettent de retrouver des bits d'informations sensibles. Plusieurs techniques statistiques sont disponibles, et le *deep learning* ou apprentissage profond est une des techniques récentes dont dispose l'attaquant. Une analyse statistique comme la DPA a fait l'objet de nombreuses publications.

La **DPA (*Differential Power Analysis*)** est connue depuis la fin des années 1990 et constitue une bonne illustration de la nécessité de considérer le matériel et le logiciel comme un couple indissociable,

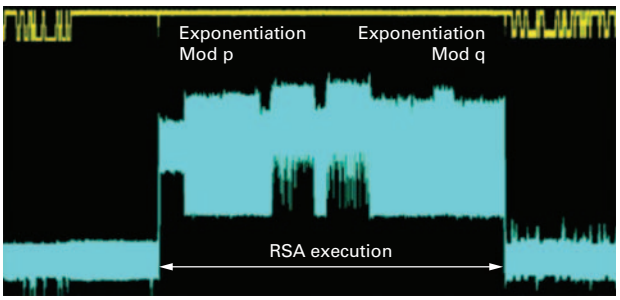


Figure 8 – Mesure effectuée lors d'un calcul cryptographique RSA

non seulement dans une carte à puces, mais également dans tout dispositif cryptographique moderne. Ces attaques sont fondées sur une analyse statistique des variations de puissance consommée pendant l'exécution d'un algorithme cryptographique connu :

– l'objectif est, par exemple, de déterminer la clé secrète utilisée dans un calcul DES, en collectant tout d'abord N traces du signal électrique S_{ij}, i étant l'indice de la trace et j celui de l'échantillon de signal considéré ;
– l'étape suivante consiste à répartir les traces en deux familles, S_0 et S_1, selon une fonction de partition P telle que :

$$S_a = \left\{ S_{ij} \times P(.,.,.) = a \right\}$$

avec a = 0,1 ;
– puis de calculer la moyenne du signal de puissance pour chaque famille :

$$A_a[j] = 1/S_a \; \Sigma S_{ij} \quad \text{pour } S_{ij} \; \hat{I} \; S_a$$

avec $\Sigma S_a = N$.

Pour la fonction de partition, on fait, par exemple, le choix d'un bit intervenant en un point particulier du calcul cryptographique (comme une sortie d'une boîte S du DES), en supposant que l'on connaisse une partie de la clé secrète (par exemple, les 6 bits de la clé secondaire intervenant à l'entrée de la boîte S). Une légère différence de puissance consommée apparaît selon que ce bit est égal à 0 ou 1. En considérant toutes les valeurs possibles de la clé partielle (ici 64) pour un grand nombre d'échantillons et en faisant la différence des puissances moyennes consommées respectives, on accroît l'amplitude de cette différence, à tel point que sa représentation graphique fait apparaître des « pics DPA » lorsque la clé est bonne, tandis que le signal tend vers un bruit blanc pour toutes les autres valeurs (voir figure **9**) ;
– en répétant l'opération sur les autres boîtes S, un attaquant peut obtenir séquentiellement les 48 bits des clés secondaires, le dernier octet étant ensuite facilement obtenu.

Bien d'autres variantes de la DPA existent, en utilisant les mêmes principes de base, y compris sur d'autres algorithmes que le DES. Cet exemple montre à quel point le matériel et le logiciel sont intimement liés dans les menaces et dans les contre-mesures à mettre en place. On voit aussi à travers cet exemple la nécessité de réunir des compétences en électronique, en traitement du signal et en cryptographie.

Les attaques par perturbation consistent à modifier l'exécution du programme embarqué dans le dispositif. Lorsque celui-ci est en exécution, une perturbation peut en effet engendrer une « faute » du microcontrôleur, modifier l'exécution normale du programme et conduire à un comportement avantageux pour l'attaquant. Une faute peut par exemple autoriser l'accès à une donnée sensible inaccessible autrement. Elle peut également générer un calcul cryptographique erroné, mais exploitable par l'attaquant : c'est typiquement le principe des *Differential Fault Analysis* (DFA), dont la plus célèbre est l'attaque de Bellcore contre l'algorithme RSA. Il existe plusieurs

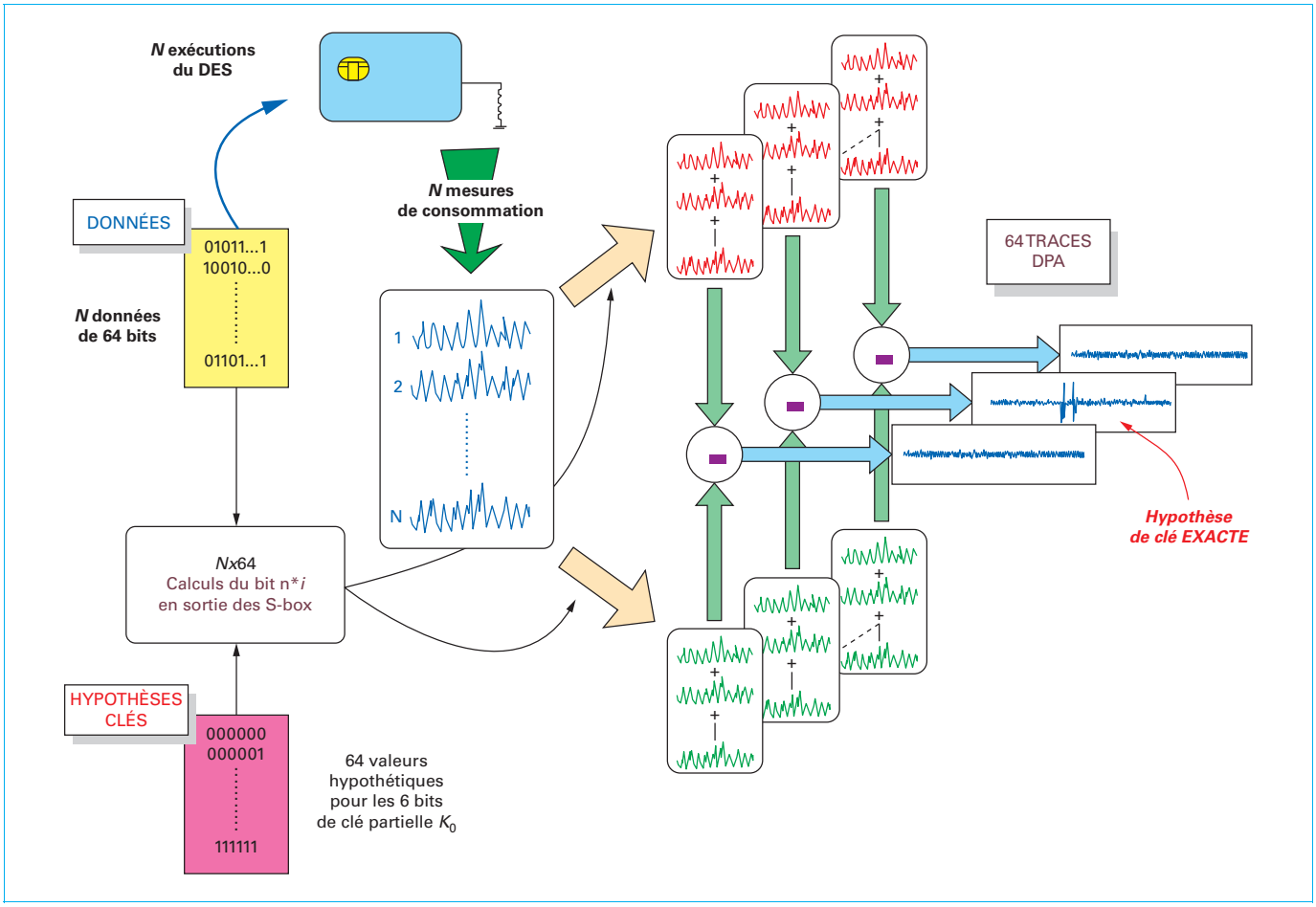


Figure 9 – Traces DPA lors de l'exécution d'un calcul DES

façons de provoquer une perturbation, la plus simple étant la génération de *glitch* sur les signaux d'entrée/sortie, c'est-à-dire une variation brève mais significative du signal. La plus élaborée consiste à injecter une impulsion laser sur un élément spécifique du circuit en fonctionnement.

■ **Les attaques dites intrusives** modifient physiquement le dispositif afin d'en extraire une donnée sensible. Là encore, un attaquant dispose de plusieurs types d'outils pour pénétrer un circuit électronique. Le plus souvent, il utilise des outils employés par l'industrie du semi-conducteur à des fins de diagnostic d'erreur ou de correction de circuits. Ses outils sont en fait détournés de leur usage originel, parfois modifiés, pour une utilisation malveillante.

3.2.2 Protections

Les cartes à puce ont la particularité d'être protégées contre l'ensemble des attaques décrites plus haut. Une description rapide des contre-mesures mises en œuvre est donnée ci-après.

■ **Protections contre les attaques logiques** : par construction, la surface d'attaques logiques contre une carte à puce est étroite. En effet, une carte à puce offre une interface logique limitée, à travers un jeu de commandes réduit. Les commandes sensibles sont protégées par des mécanismes cryptographiques éprouvés. La simplicité du jeu de commandes permet au développeur du logiciel embarqué de coder le traitement de chaque commande avec précaution et d'éviter les erreurs de programmation conduisant par exemple à des débordements de piles, de tableaux, etc. La simplicité du jeu de commandes permet également une validation fonctionnelle proche de l'exhaustivité. Les cartes dites « ouvertes » autorisent le chargement à posteriori de nouvelles applications. Ce chargement se fait alors via des protocoles standardisés comme ceux définis par GlobalPlatform. Enfin, une part significative des cartes à puce repose sur la technologie JavaCard, qui offre par essence de solides propriétés d'isolation.

■ **Protections contre les attaques par canaux cachés** : les algorithmes cryptographiques sont mis en œuvre avec un jeu de contre-mesures élaborées, destinées à rendre l'exploitation des mesures effectuées difficile voire impossible : les paramètres des algorithmes sont randomisés d'une exécution à l'autre, l'ordre des calculs varie également d'une exécution à l'autre, l'algorithme s'exécute en temps constant, voire à code constant si nécessaire… Le microcontrôleur injecte aléatoirement des instructions additionnelles, l'horloge du circuit varie aléatoirement…

■ **Protections contre les attaques par perturbation** : les protections consistent le plus souvent à détecter un comportement erroné du programme. Pour ce faire, les processus sensibles sont redondés afin de comparer leur résultat. S'ils diffèrent, c'est qu'une erreur s'est produite dans l'un d'entre eux et le dispositif prend alors les mesures nécessaires. Le dispositif dispose également d'un ensemble de capteurs qui détectent un environnement d'exécution hostile. Le circuit est lui-même composé de structures redondées et

comparées systématiquement. Les mémoires disposent de mécanismes assurant leur intégrité.

■ **Protections contre les attaques intrusives** : les mémoires embarquées dans le dispositif sont chiffrées et brouillées. Les bus transportant les bits d'information entre les mémoires et le microprocesseur sont chiffrées. Le circuit peut le cas échéant être recouvert d'un niveau de métal « bouclier ». La logique du circuit est mélangée et saupoudrée de capteurs.

3.2.3 Fabrication et personnalisation

La sécurité intrinsèque d'un dispositif est nécessaire, mais pas suffisante, pour assurer un bon niveau de sécurité à l'application qui l'utilise. En effet, comme dans tout système, il faut que tous les maillons de la chaîne respectent certaines règles. En particulier, les procédures de fabrication et de personnalisation des dispositifs sensibles comme les cartes à puces constituent des étapes critiques. Des clés cryptographiques doivent en effet être injectées dès la fabrication du circuit. Des clés diversifiées doivent ensuite être générées, puis injectées dans le logiciel embarqué par l'intermédiaire d'une carte mère hautement sécurisée, le *Hardware Security Module* (HSM). La gestion des HSMs est elle-même un point critique de sécurité dans la chaîne de fabrication et de personnalisation.

3.3 Certification

Une certification de sécurité est une **évaluation sécuritaire d'un produit ou d'une solution** conduite par une tierce partie reconnue et indépendante du vendeur. Cette tierce partie s'appuie en général sur des laboratoires spécialisés et indépendants pour d'une part réaliser les tests de sécurité et, pour d'autre part mettre en œuvre les audits des organisations et des processus concernés. Lorsque le laboratoire mandaté conclut à la résistance du produit ou de la solution conformément aux exigences attendues, et que les audits confirment que le cycle de vie du produit jusqu'à sa distribution respecte les règles de sécurité, la tierce partie atteste des résultats à travers la délivrance d'un certificat.

La grande majorité des applications utilisant une carte à puces exige une certification préalable de la carte et de ses processus de conception et de fabrication. Depuis son origine, la carte à puces fait ainsi l'objet de certifications, qui sont autant de contrôles de la solidité des produits déployés.

Trois grandes familles de certification de sécurité s'appliquent aux cartes à puces :

– certifications par les agences nationales de sécurité du SOG-IS,
– certifications par les schémas de paiement : EMVCo, Visa, Mastercard, AMEX…,
– certifications FIPS.

Le SOG-IS est un accord de coopération et de reconnaissance mutuelle entre les principales agences nationales de sécurité européennes : BSI (Allemagne), ANSSI (France), NLNCSA (Pays-Bas), CCN (Espagne)… La certification des cartes à puces par le SOG-IS se fait suivant la méthodologie des critères communs, dans une version adaptée aux cartes à puces. La résistance de la carte aux attaques mentionnées dans le paragraphe précédent et la conformité des processus de développement et de fabrication sont prises en compte dans les certifications du SOG-IS. Suite à l'adoption du Cybersecurity Act le 7 juin 2019, la certification des cartes à puces entre dans la catégorie des niveaux « élevés » introduits par le Cyberact. Le SOG-IS est logiquement amené à évoluer vers un cadre européen, le *European Cybersecurity Certification Group*, dont une des premières missions est la mise en place, en collaboration avec l'ENISA, l'Agence européenne de sécurité, d'un schéma de certification *Common Criteria Scheme* pour les cartes à puces et autre dispositifs de haute sécurité.

Les schémas de paiement (Visa, Mastercard, AMEX…) ont mis en place leurs propres schémas de certification qui reposent sur un socle commun défini par EMVCo. La résistance de la carte aux attaques mentionnées dans le paragraphe précédent et la conformité des processus de fabrication sont prises en compte lors de ces certifications. Les schémas de paiement reconnaissent les certifications conduites par les membres du SOG-IS.

Les certifications FIPS sont gérées par le NIST (États-Unis), et pour les cartes, elles se concentrent sur la mise en œuvre des algorithmes cryptographiques. Seules la conception et la validation fonctionnelle des algorithmes cryptographiques sont prises en compte dans les certifications FIPS.

Une des règles de base en sécurité de l'information est la **nécessité de faire valider une solution par des experts indépendants**. Et, pour ces experts, la validation d'un produit ou d'une solution doit passer par une tentative pratique de les « casser », les experts jouant le rôle des hackers utilisent tous les moyens pour mettre à mal le produit ou la solution à valider : c'est ce qu'on appelle communément le *ethical hacking*. Les attaques n'ont ici pas d'objectifs malveillants, mais au contraire la vertu d'éprouver une solution, pour éventuellement la renforcer avant qu'elle ne soit mise à la disposition des utilisateurs.

Depuis l'origine, les acteurs de la carte à puce ont mis en place des évaluations sécuritaires basées sur **un catalogue d'attaques systématiquement mises en œuvre**. Ainsi, l'ensemble des attaques décrites dans les paragraphes précédents sont appliquées lors des évaluations conduites par les laboratoires indépendants du SOG-IS ou des schémas de paiement. Pour que ces laboratoires travaillent sur une base commune, pour qu'ils appliquent des attaques d'un niveau comparable, et pour qu'ils mettent en œuvre une panoplie complète de tests de sécurité, un catalogue d'attaques est tenu à jour depuis le milieu des années 2000 par un groupe d'experts. Le groupe JHAS, formé d'experts issus de l'industrie, des laboratoires indépendants et des organismes de certification (ANSSI, BSI, EMVCo…), se réunit tous les 2 mois pour discuter des nouvelles attaques à expérimenter lors des évaluations de cartes. Le travail de ce groupe, à travers le catalogue d'attaques, est la pierre angulaire de la sécurité des cartes à puces et de leurs certifications. C'est sur les travaux du JHAS que reposent les certifications par les agences nationales de sécurité du SOG-IS et par les schémas de paiement EMVCo, Visa, Mastercard, AMEX…

Cet ensemble de schémas de certification sécuritaires fait de la carte à puces la technologie la plus contrôlée du monde de l'information.

4. Construction

4.1 Principes de construction

Une carte à puces est constituée de trois éléments (figure **10**) :

– une carte en matière plastique (le plus souvent en PVC-chlorure de polyvinyle, ABS-acrylonitrile-budadiène-styrène ou PC – polycarbonate), possédant ou non une piste magnétique ;
– un module électronique, supportant les contacts électriques ou un transpondeur (module électronique et antenne) ;
– un circuit intégré en silicium.

Pour de plus en plus d'applications, le circuit intégré possède une **interface sans contact**, utilisant les mêmes principes que l'identification par radiofréquence (RFID). Dans ce cas, on utilise le principe de type transpondeur magnétique. Un lecteur envoie un signal radio à la carte qui, elle-même, contient une antenne déposée sur le substrat en plastique et connectée au circuit intégré. La puce est alors alimentée par le signal radio et va ainsi pouvoir communiquer avec le lecteur. Les intensités du champ magnétique ainsi créées sont comprises entre 1,5 et 7,5 $A.m^{-1}$, la distance de fonctionnement entre le lecteur et la carte varie de 0 à 10 cm, la fréquence utilisée étant normalisée à 13,56 MHz. Cette technologie du sans contact trouve de nombreux champs d'applications, tels que

le paiement (y compris sur mobile), la billettique, les cartes d'identité, le contrôle d'accès, les passeports électroniques, etc. Le changement de format, du format carte de crédit au format passeport, est un des bénéfices du « sans contact » qui ne nécessite plus un placement précis du produit dans un connecteur.

Les composants sont réalisés par les fabricants de semiconducteurs à partir de tranches circulaires de silicium en utilisant une série de masques pour réaliser la photolithogravure des circuits. Une même tranche (ou *wafer*) de 8 ou 12 pouces de diamètre peut ainsi contenir plusieurs milliers de puces rectangulaires, qu'il faut découper à l'aide de scies diamantées (figure **2**). Des machines automatiques viennent ensuite prélever ces puces pour les fixer au dos des modules, dans une cavité qui a été ménagée à cet effet, c'est l'opération dite de *die bonding*.

4.2 Interconnexion des composants

Cette opération consiste à **relier électriquement les plots du composant aux contacts électriques du module**. Ce câblage peut s'effectuer selon plusieurs techniques de base : le *wire bonding* se fait en utilisant du fil d'or ou d'aluminium, dont le diamètre est voisin de 10 à 20 µm, et dont les soudures sur les plots des composants sont réalisées par thermocompression ou ultrasons. Outre la fragilité relative des fils, cette technique, qui est la plus répandue, nécessite la répétition des soudures pour chaque extrémité des fils.

Une autre technique très fiable appelée TAB (*Tape Automated Bonding*) utilise **un ruban continu découpé et soudé directement sur les composants** en une seule opération par l'intermédiaire d'un métal d'apport. Il est alors possible de tester les composants unitairement sur le ruban avant leur montage. Cette technique a permis de réaliser les toutes premières cartes, mais elle a souvent cédé la place à la précédente pour des raisons de coût et de flexibilité.

On peut également réaliser l'interconnexion en utilisant **une grille métallique semblable à celle des composants classiques avec un raccordement en fils**, mais ce procédé nécessite une industrialisation poussée pour satisfaire aux contraintes de durabilité et d'encartage. Les modules ainsi câblés sont protégés par une résine d'enrobage avant d'être encartés.

Le lecteur pourra aussi consulter l'article [E 3 401].

4.3 Encartage

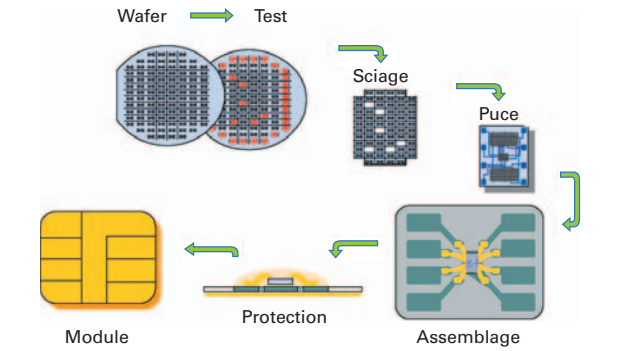
L'encartage consiste à fixer par collage les modules électroniques dans la carte, selon des procédés de fabrication permettant d'obtenir la qualité et la fiabilité requises. La carte terminée doit en effet résister aux agressions mécaniques et climatiques très sévères auxquelles elle est soumise dans les utilisations courantes.

Les modules sont **encartés à l'aide de chaînes industrielles robotisées**. Lorsque la carte comporte une piste magnétique, toutes ces opérations doivent être particulièrement maîtrisées afin de ne pas perturber le matériau et de conserver une parfaite planéité de la surface. La figure **11** donne la coupe transversale d'une carte à puces montrant les différents éléments. Il faut noter que le module doit tenir dans une épaisseur inférieure à 0,6 mm, compte tenu de l'épaisseur normalisée des cartes qui est de 0,76 mm. Les dernières technologies disponibles ont permis de diminuer cette épaisseur de module jusqu'aux environs de 150 µm, ce qui permet éventuellement de réaliser des structures verticales par empilement de puces (par exemple, un microcalculateur et une mémoire).

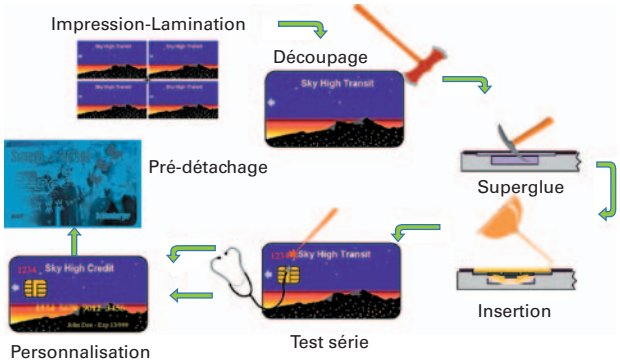
Lorsque la carte est terminée, **un test automatique** vient vérifier que le composant fonctionne correctement avant d'être personnalisé. Cette dernière étape comprend deux phases distinctes :

– la **personnalisation graphique** consiste généralement à embosser, à imprimer au jet d'encre ou à graver au laser des informations propres à l'utilisateur dans le plastique ;

– la **personnalisation électrique** consiste à encoder, s'il y a lieu, la piste magnétique et à inscrire les droits, les caractéristiques d'utilisation, les éléments de sécurité (clés, certificats…) dans la mémoire de la carte. Cette dernière opération devant s'effectuer de façon sécuritaire pour des raisons évidentes, elle est généralement réalisée après le contrôle de codes ou de clés de personnalisation, eux-mêmes calculés par des cartes mères qui sont capables de recalculer les clés diversifiées des cartes filles. En effet, chaque carte possède son propre jeu de clés pour protéger chaque composant pendant toutes les phases de la fabrication des cartes. Si un composant détecte une tentative d'intrusion non licite à une phase déterminée, il se bloque définitivement.

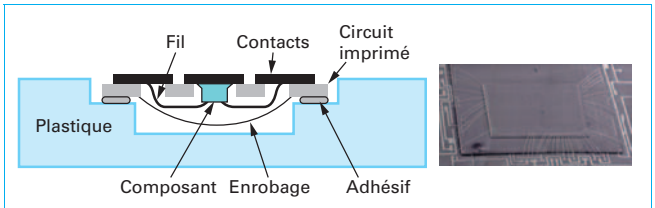


(a) **asemblage de la puce**



(b) **encartage du micromodule**

Figure 10 – Fabrication d'une carte à puces



(a) **coupe transversale d'une carte à puces standard**

(b) **example par empilement de puces**

Figure 11 – Encartage de modules

4.4 Connectique

Le module électronique supporte les contacts qui assurent les différentes liaisons électriques avec le monde extérieur. La métallurgie, l'état de surface et la topologie des contacts sont particulièrement délicats pour assurer une connectique fiable. Il est important de noter que la qualité des contacts est étroitement liée à la conception du connecteur, utilisé dans les machines de traitement. Il existe maintenant sur le marché toute une gamme de connecteurs de technologie et de qualité très variables. Un connecteur à bas coût ne dépassera pas 10 000 manœuvres, tandis que les meilleurs modèles peuvent dépasser plusieurs centaines de milliers de manœuvres dans des conditions difficiles. On aura soin de distinguer entre les connecteurs « frottant », qui peuvent entraîner des actions abrasives néfastes, et les connecteurs à contacts « atterrissant », qui ont une fiabilité nettement supérieure sans endommager les cartes.

4.5 Évolution vers le « sans contact »

De plus en plus d'applications souhaitent offrir la flexibilité, la facilité d'utilisation et la rapidité de la solution radio fréquence « sans contact ». La figure **12** illustre les différentes variantes de cartes à puces.

Nous distinguerons 3 types de produit « sans contact » :

– la carte *dual interface*,
– la carte « sans contact »,
– la carte hybride.

4.5.1 Carte *dual interface*

La carte *dual interface* encore appelé carte Combi peut être utilisée en mode contact et en mode sans contact, elle ne possède qu'une puce qui satisfait aux deux protocoles.

Sa construction repose sur l'utilisation d'un *inlay antenne*, généralement une couche de plastique en PVC ou polycarbonate comprenant une antenne en cuivre ou aluminium gravée. Cet ensemble est inséré entre d'autres feuilles plastiques, avant que l'ensemble soit laminé en pression et température, puis découpé au format carte.

Lors de l'opération d'encartage du module, les plots du module sont connectés par colle argent ou matériau conducteur polymère à l'antenne.

4.5.2 Carte sans contact

La carte sans contact à la différence de la carte *dual interface* ne possède qu'une interface sans contact.

Sa construction repose complètement sur l'utilisation d'un *inlay sans contact* légèrement plus sophistiqué que l'*inlay antenne* de la carte *dual interface*. Pour cette construction, l'antenne gravée ou insérée par ultrason est directement connectée avec le module par thermocompression ou au moyen d'une colle ACF (*Anisotropic Conductive Film*). Il reste à assembler les différentes couches de plastique, à laminer et à découper pour obtenir la carte prête à l'emploi.

Il est à noter que la majorité des cartes sans contact utilisent un module pour protéger la puce en silicium et permettre une connexion robuste avec l'antenne. Le module et la puce sont cependant plus fins – de l'ordre de 50 à 75 µm pour la puce – pour loger à l'intérieur de la carte à puces.

Quelques industriels de la carte reportent la puce nue en *flip chip* sur l'antenne, réalisée par gravure chimique ou par sérigraphie d'une encre argent.

4.5.3 Carte hybride

Comme l'indique son nom, la carte hybride est une carte contact et sans contact qui à la différence de la carte *dual interface*, utilise deux puces : une avec module contact et une autre directement connectée à l'antenne. On part donc d'une carte sans contact puis on usine une cavité pour loger le module contact.

4.6 Contraintes particulières des documents d'identité gouvernementale

Les cartes d'identité ou les passeports numériques se distinguent des autres cartes à puces par l'importance du support physique pour garantir une longue durée de vie et une sécurité physique à l'identique d'un billet de banque. Les documents d'identité sont l'objet de tentative de fraude pour commettre des actes délictueux allant du vol d'identité, au terrorisme et au trafic humain.

Pour garantir une durée de vie de dix ans et plus, les plastiques traditionnelles comme le PVC (polychlorure de vinyle), ou l'ABS (acrylonitrile butadiène styrène) ne sont plus utilisés et doivent être remplacés par des alliages PVC/PET (polytéréphtalate d'éthylène) ou par du polycarbonate (PC).

Le polycarbonate s'est progressivement imposé comme le matériau de choix pour assurer la longévité, mais aussi pour permettre d'inclure en sous-couche des éléments de sécurité comme les impressions de sécurités réalisées en offset – lithographie (guillochés, fine lignes en ton direct, micro-texte, impression irisée, etc.), l'utilisation d'encres spéciales en sérigraphie et le dépôt d'hologrammes.

Les différentes feuilles de polycarbonate, constituant le document d'identité, sont agencées pour accueillir l'ensemble des éléments de sécurité imprimés, les inlays, les hologrammes ou DOVID (*Diffractive Optical Variable Image Device*), les fenêtres transparentes interdisant des copies par moyen d'impression numérique, et d'autres solutions plus ou moins sophistiquées et souvent propriétaires.

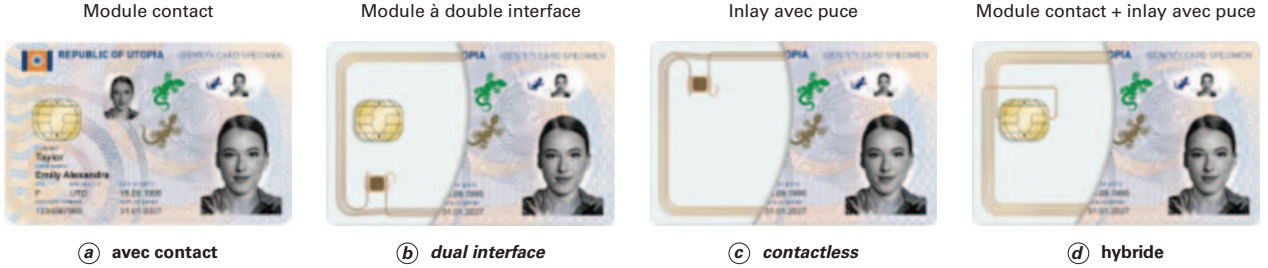
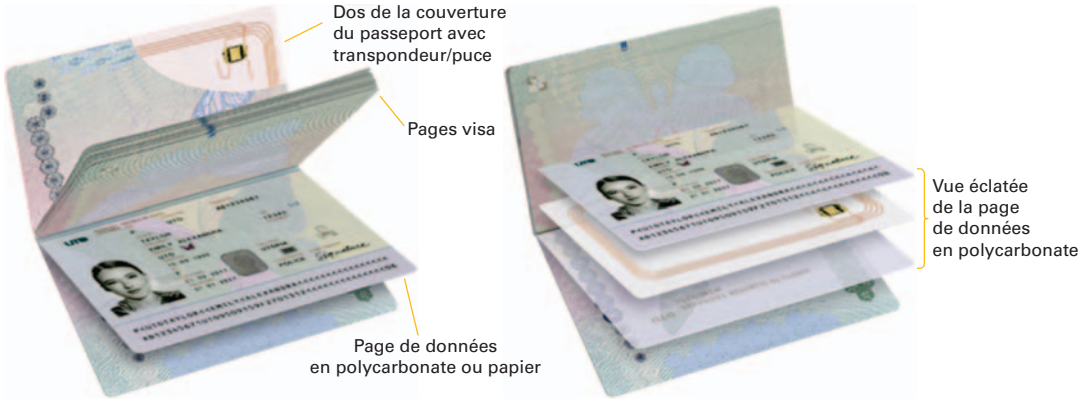


Figure 12 – Illustration des différentes solutions de cartes à puces (source Gemalto)



(a) couverture électronique avec puce (b) page de données électronique avec puce

Figure 13 – Illustration des différentes solutions de passeports électroniques (source Gemalto)

Les feuilles assemblées sont laminées sous pression et sous température pour fusionner les différentes couches de polycarbonate et encapsuler d'une manière définitive les éléments de sécurité.

La sécurité physique du document est encore renforcée lors de la personnalisation graphique par l'utilisation de Laser Infra Rouge (IR) gravant des photos en niveau de gris et du texte à l'intérieur des couches. Il existe aussi depuis peu la possibilité de marquer au laser des photos en couleur à l'intérieur du document.

La figure **13** illustre les deux types de passeports électroniques sur le marché : le passeport qui embarque la puce et l'antenne dans la couverture, et celui qui contient une page de données en matériau polycarbonate avec la puce et l'antenne directement intégrées dans le plastique.

4.7 Aperçu de l'écosystème industriel

Le monde de la carte à puces est dominé en 2019 par trois acteurs majeurs : Thales qui a acheté Gemalto en avril 2019, Idemia qui est le résultat de la fusion de Morpho et Oberthur, et la société Veridos qui est le bras armé de Giesecke & Devrient et de la Bundesdrukerei pour les marchés de l'export. Ces sociétés sont concurrencées par des entreprises plus modestes comme les Chinois Golpac, Eastcompeace et d'autres acteurs nord et sud-américains comme CPI Card Group et Valid. La concurrence ne cesse de croître grâce à l'innovation et les activités de fusion et d'acquisition dans le secteur.

La plupart de ces sociétés développent eux-mêmes les systèmes d'exploitation et produisent les cartes à puces. Ils peuvent aussi assurer la personnalisation des cartes, systématiquement pour les cartes GSM, souvent pour les cartes bancaires, et plus rarement pour les cartes d'identité et les passeports électroniques. Ces derniers documents de sécurité sont souvent personnalisés par des entités sous contrôle gouvernemental comme les imprimeries nationales.

Les puces et les semiconducteurs sont encore largement développés et produits par les sociétés de conception et fabrication de puces silicium comme Infineon, NXP, Samsung, ST Microelectronics et Inside Contactless, mais les acteurs majeurs de la carte à puces ont démarré leur propre design de puces et en sous-traitent la fabrication à des fonderies de silicium indépendantes comme TSMC, UMC ou GlobalFoundries.

Avec la croissance des applications d'identité, des transferts de technologies des différents acteurs de la cartes à puces vers les imprimeries nationales se font jour. Elles sont poussées par les gouvernements pour augmenter la valeur ajoutée dans le pays et renforcer la sécurité, mais de fait augmentent la concurrence entre les différents acteurs traditionnelles et réduisent le marché accessible.

5. Systèmes d'exploitation

5.1 Explication globale et mécanismes de base

5.1.1 Fonctionnalités

L'usage courant a consacré le terme de **système d'exploitation** (***Operating System***, **OS**) pour désigner le logiciel de commande d'une carte à microcalculateur qui interprète et exécute les différents ordres élémentaires que cette carte peut réaliser. Situé dans la partie mémoire persistante du microcalculateur (ROM, EEPROM ou FLASH), il est implanté pour la ROM dans l'un des masques qui sert à la fabrication du circuit intégré. Ainsi, par abus de langage, les notions d'OS et de masque sont couramment confondues.

L'OS réalise principalement les fonctions suivantes :

– gestion des échanges entre la carte et le monde extérieur, notamment le protocole d'échanges,
– gestion des différents fichiers et des données à l'intérieur de la mémoire,
– contrôle des accès aux informations et aux fonctions (par exemple : sélection de fichier, lecture, écriture, modification de données),
– gestion de la sécurité de la carte et de la mise en œuvre des algorithmes cryptographiques,
– fiabilité du fonctionnement, notamment cohérence et intégrité de certaines données, ruptures de séquences et reprise des erreurs,
– gestion du cycle de vie de la carte dans ses différentes phases (fabrication, personnalisation, utilisation, fin de vie).

Un système d'exploitation de carte à microcalculateur est **similaire aux systèmes utilisés dans les micro-ordinateurs**. Il est l'une des parties les plus importantes des cartes à microcalculateur, car c'est lui qui leur confère toutes leurs fonctions, en particulier vis-à-vis de la

supervision des fonctions sécuritaires. Les bons systèmes d'exploitation sont très complexes à réaliser, et leur implémentation doit tenir compte à la fois du faible coût, de la sécurité, de l'architecture de la machine, des performances, de la compacité, de la standardisation et de la fiabilité. Le système d'exploitation et le composant ne font qu'un, ils constituent un couple indissociable, et cela est une démarche inhabituelle, pour qui développe du logiciel classique.

5.1.2 Architectures

Les OS de cartes se divisent en deux familles :

– d'une part, les **OS fermés**, souvent liés à une application spécifique et qui possèdent des commandes adaptées à cette application, allant parfois jusqu'à la contenir elle-même ;

– d'autre part les **systèmes ouverts**, autorisant le téléchargement (sécurisé) dans la carte de logiciels ou d'applications après sa délivrance à l'utilisateur.

Les OS pour cartes à puces sont maintenant logés en mémoire flash. Il reste très peu de composants offrant de la ROM. Historiquement écrits en assembleur, optimisés pour une unique application, ils ont considérablement évolué depuis les vingt dernières années et intégré la plupart des concepts des OS traditionnels, à l'exception notable, pour des raisons de sécurité, d'une interface utilisateur. Beaucoup d'entre eux supportent le concept de multi application. Pour des raisons d'efficacité, les OS modernes sont maintenant principalement écrits en langage de haut niveau (C, C++, Java), présentent un caractère en couches, avec seulement les routines de bas niveau (celles liées à l'interface avec le hardware) écrites en Assembleur ou C. L'accroissement de la taille du code généré par la programmation de haut niveau est évidemment largement compensé par l'évolution des tailles de mémoire disponible dans les micro-processeurs. De plus, l'écriture d'une large partie du code en langage de haut niveau est un gage important de portabilité du logiciel d'un composant à un autre (ce qui constitue, souvent, une exigence forte des utilisateurs). La figure **14** présente l'architecture classique d'un OS moderne en couches pour cartes à puces. La couche de services génériques inclut en particulier un gestionnaire d'APDU, qui joue un rôle particulièrement important et permet l'interfaçage avec toutes les applications (ce qui n'était pas le cas des OS spécifiques fermés sur un applicatif déterminé).

Pour assurer la sécurité des accès et des fonctions, la couche générique fait appel à des algorithmes cryptographiques symétriques, mais aussi asymétriques, souvent associés à de nouveaux composants. Ces derniers intègrent des coprocesseurs, comportant à leur tour des opérateurs câblés, tels que des multiplications modulaires rapides ou encore des microcalculateurs à architecture RISC (*Reduced Instruction Set Computer*), ou DSP (*Digital Signal Processor*).

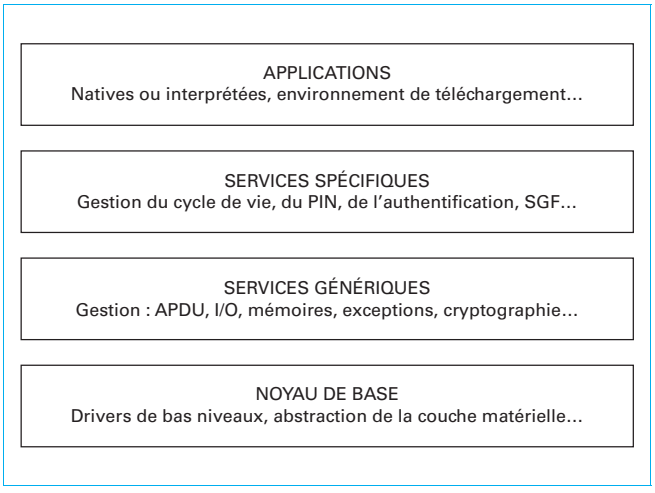


Figure 14 – Architecture en couche d'un OS moderne pour carte à puces

Les fonctions d'authentification et de signature électronique prennent de plus en plus d'importance dans les OS. C'est le cas en particulier pour les diverses cartes à usage gouvernemental (cartes d'identité électronique, passeports, etc.).

Outre les fonctions élémentaires permettant d'accéder à la carte (lectures et mises à jour), un certain nombre de mécanismes de base assurent la sécurité en s'appuyant sur une structure sous-jacente des informations.

Pour l'organisation des fichiers, la plupart des OS de cartes se rapprochent des recommandations de l'ISO qui a normalisé l'architecture arborescente déjà décrite en cartes à puces.

Cette organisation, qui peut sembler familière, cache en réalité **des aspects importants**. La sécurité des accès y est omniprésente à travers les attributs de sécurité affectés à chaque fichier, qui fixent les conditions devant être satisfaites pour exécuter les opérations demandées. Un certain nombre de droits doivent être acquis pour l'utilisation normale de la carte, la création de nouveaux fichiers ou la modification éventuelle de leurs attributs. Ainsi, l'OS protège les différents fichiers en fonction du niveau où ils se trouvent. Il assure l'indépendance des applications et gère la transmission héréditaire des droits lorsqu'on accède aux fichiers à l'intérieur de la carte.

Tous les OS pour cartes à puces (fermés comme ouverts) doivent garantir l'atomicité d'une transaction, c'est-à-dire le fait que cette dernière s'est exécutée intégralement (et correctement) ou pas du tout. Aucun état intermédiaire, générateur d'une situation instable du point de vue de la sécurité, n'est toléré (cela est particulièrement important pour les transactions bancaires). À ce niveau, et avec les spécificités propres à l'espace réduit disponible en mémoire, il est donc nécessaire d'implémenter des fonctions de même nature que celles réalisées pour garantir l'intégrité des bases de données ou des mémoires secondaires des ordinateurs. Sans entrer ici dans les détails, ces mécanismes sont basés sur l'utilisation de buffers auxiliaires en EEPROM ou en flash, qui, par des mécanismes de flags, sont validés puis invalidés au fur et à mesure de l'avancement des opérations (écriture, lecture, mise à jour, effacement) sur les fichiers.

5.1.3 Mécanismes de base

Parmi les mécanismes de base d'un OS de carte, trois d'entre eux sont absolument fondamentaux :

– **l'authentification dynamique** est réalisée à partir d'un dialogue « Question-Réponse » non reproductible avec la carte qui permet de prouver :

- que la carte ou un fichier de cette carte ont bien été émis par un organisme habilité,
- que la carte n'a pas été falsifiée ou contrefaite ;

Pour être sûre, l'authentification dynamique doit mettre en œuvre un dialogue aléatoire avec le monde extérieur, de façon à éviter toute simulation et re-jeux. Des méthodes d'authentification ont déjà été décrites au paragraphe 3. Une carte peut authentifier un terminal par un processus identique dans l'autre sens ; on réalise ainsi une **authentification mutuelle**. Un protocole à apport nul de connaissance peut également être utilisé (algorithmes asymétriques) ;

– **la signature électronique**, objet d'une directive de l'Union européenne transposée dans les législations des membres de l'UE, permet, suivant une norme ISO/IEC, d'authentifier des messages en vérifiant leur origine et leur intégrité. Les schémas de signature sont surtout fondés sur des algorithmes asymétriques. Les fichiers à signer sont en général transformés par une fonction de « hachage », avant la signature proprement dite. Possédant la clé publique du signataire, n'importe qui peut la vérifier – mais aussi s'assurer que le contenu du fichier n'a pas été altéré ou modifié pendant la transmission ;

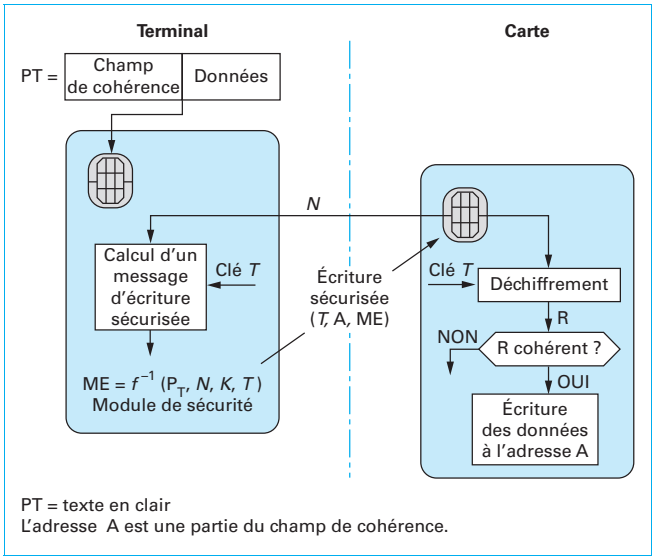
– **la certification des données** de la carte est une extension de la notion de signature électronique, qui prouve qu'une information se trouve effectivement à l'intérieur d'une carte donnée. Les architectures correspondantes, appelées PKI (*Public Key Infrastructure*) ou IGC (Infrastructure de gestion de clés) s'appuient également sur un cryptosystème à clés publiques. Chaque utilisateur du système possède un certificat généré et géré par des autorités indépendantes (au moins une autorité d'enregistrement et une autorité émettrice) contenant, entre autres données, la clé publique de l'utilisateur. Le certificat original est souvent stocké dans une carte à puces, et permet de générer à la volée, par la carte, des informations confidentielles (*via* sa clé privée) liées à une transaction. Ces données constituent en fait une signature numérique, et elles sont immergées de manière cryptographique avec d'autres données propres à la transaction et transmises au centre de traitement des transactions. Le serveur distant recalcule la signature à partir des données de transaction, et la compare avec celle qu'il peut construire à partir des données publiques de l'utilisateur. L'identité des deux signatures ainsi générées est une preuve d'opération licite, car seule la carte contenant le bon certificat aura pu engendrer un tel résultat.

5.1.4 Ecriture sécurisée

La troisième fonction fondamentale d'un OS de carte concerne une partie très importante de la vie d'une carte, mais qui est souvent complètement mal connue des utilisateurs : comment une carte est-elle née ? comment une application a-t-elle pu s'introduire dans une carte en toute sécurité ? comment les clés cryptographiques peuvent-elles être créées ou modifiées dans une carte sans être compromises ?

L'**écriture** (ou la **mise à jour**) **sécurisée** répond à ce besoin essentiel, qui permet d'assurer le cycle de vie de la carte en toute sécurité. L'écriture sécurisée permet de transférer des données chiffrées à l'intérieur de la carte en authentifiant en même temps l'expéditeur. Ce processus est illustré sur la figure **15**.

Cette fois, c'est la carte qui émet un nombre aléatoire *N*, sur la base duquel l'organisme habilité chiffre le message PT possédant une règle de syntaxe particulière. La carte déchiffre le message et vérifie la cohérence du résultat obtenu. Si le résultat est correct, la carte inscrit l'information dans sa mémoire. Une écriture sécurisée peut également se réaliser à l'aide d'une signature électronique vérifiée au préalable.



PT = texte en clair
L'adresse A est une partie du champ de cohérence.

Figure 15 – Écriture sécurisée

5.2 Systèmes d'exploitation fermés

Il existe encore actuellement un grand nombre d'OS généraux fermés. Ces systèmes, rappelons-le, n'autorisent pas le chargement de code ou d'applications dans la carte après délivrance à l'utilisateur. Parmi ces OS, on distingue principalement deux catégories :

– les **OS monoprestataires** à cryptographie symétrique et/ou asymétrique,
– les **OS multiprestataires** à cryptographie symétrique et/ou asymétrique.

Un **prestataire** est l'entité responsable de la gestion des ressources de la carte et de la distribution des clés d'accès pour un ou plusieurs services. Il peut y avoir un ou plusieurs prestataires.

Un OS monoprestataire peut supporter en général plusieurs services, alors qu'un OS multiprestataire permet la cohabitation de plusieurs applications (chacune associée à un prestataire différent) dans la même carte. Chaque application peut à son tour abriter plusieurs services.

5.2.1 Principe des systèmes fermés

La boucle de commande d'un tel OS est décrite par le schéma synoptique de la figure **16** :

– la carte reçoit une commande (immergée dans un APDU (*Application Protocol Data Unit*)) par l'interface E/S série ;
– le gestionnaire d'E/S effectue les opérations de détection/correction d'erreurs éventuelles. Il passe le résultat au gestionnaire de message sécurisé, qui vérifie l'intégrité du message reçu ou en assure le déchiffrement, puis qui passe le résultat à l'interpréteur de commande. Cette étape est transparente si la transmission est prédéfinie comme non sécurisée ;
– l'interpréteur de commande décode l'information reçue. En cas d'impossibilité ou d'erreur, il retourne un code d'erreur approprié au terminal, *via* l'interface E/S ;
– si le décodage est correct, le gestionnaire de canaux logiques détermine le canal sélectionné, effectue le basculement du canal vers l'état logique approprié du canal et invoque la machine d'état du système, si aucune erreur n'est détectée ;
– en fonction de la commande reçue et de ses paramètres associés, la machine d'état vérifie que l'exécution de la commande est permise en fonction de l'état actuel du système : si oui, le code du programme associé à l'exécution de la commande est activé. Sinon, là encore, un code d'erreur approprié est envoyé au terminal, et la carte est remise dans un état cohérent ;
– si l'exécution d'une commande demande l'accès à un fichier, cette opération est déléguée au gestionnaire de fichiers. Ce dernier convertit les adresses logiques en adresses physiques dans le composant (le plus souvent en collaboration avec les propres mécanismes hardware du microprocesseur : ceux-ci sont alors entièrement responsables de toute la gestion et de la protection des accès physiques aux mémoires) ;
– le gestionnaire de retour est enfin responsable de l'émission d'un code global concernant l'exécution de la commande ;
– naturellement, si l'exécution de la commande nécessite des calculs cryptographiques, ceux-ci sont implémentés sous forme de routines spécialisées. Ces dernières sont invoquées soit par le gestionnaire de message sécurisé, soit par l'interpréteur de commande.

5.2.2 Exemple de systèmes fermés

Basés sur cette catégorie d'OS, on peut citer à titre d'exemple (et aussi en fonction de leur formidable succès commercial, matérialisé par des populations de cartes à plusieurs milliards d'exemplaires) :

– la version initiale de la carte SIM du radiotéléphone cellulaire européen GSM (Groupe spécial systèmes mobiles). La grande majorité des cartes SIM sont sur une architecture ouverte construite autour du standard JavaCard (§ 5.3.2), qui leur confère une structure de système multiprestataire ;
– la carte bancaire française qui suit le standard EMV ;

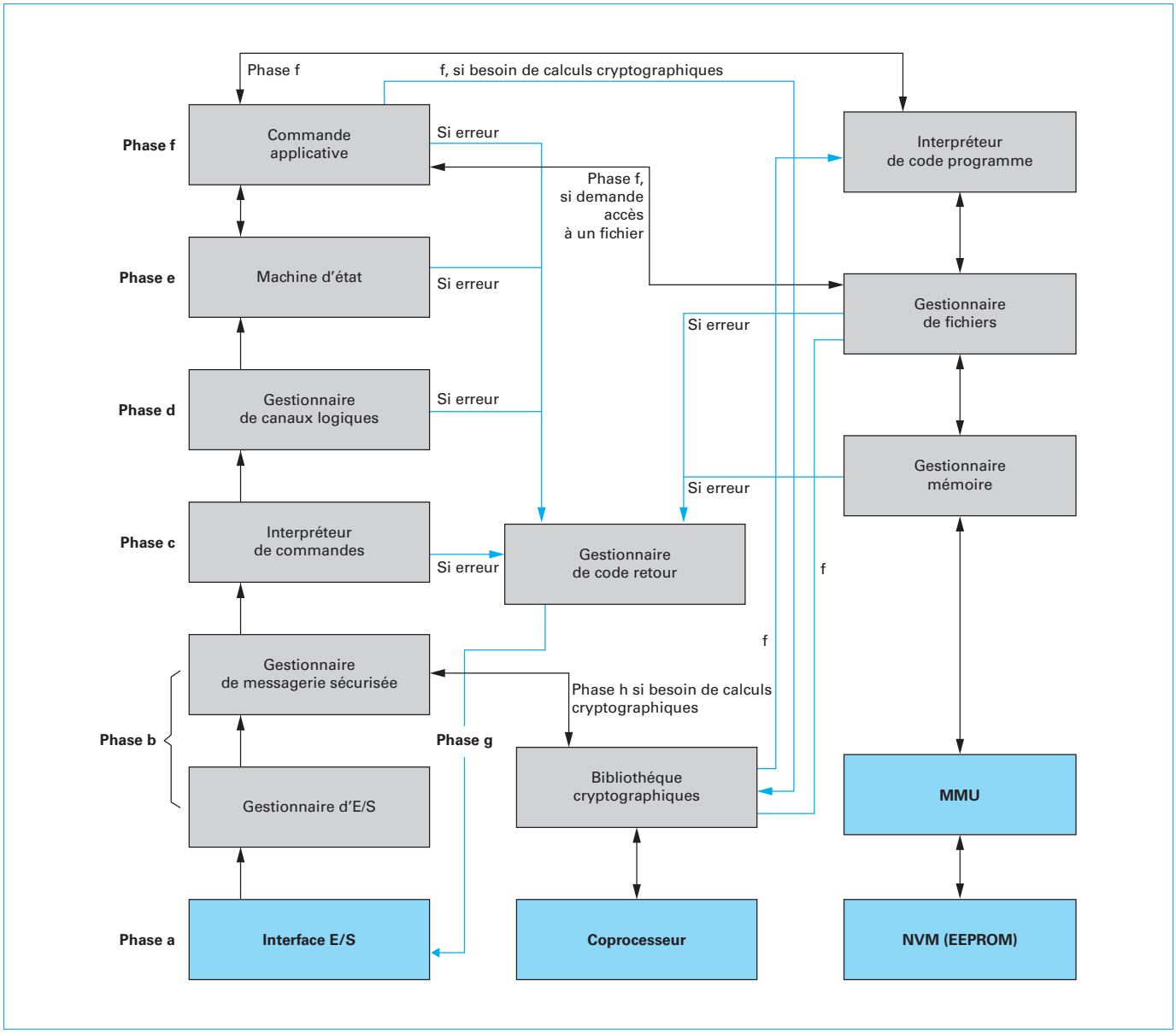


Figure 16 – Boucle de commande d'un OS pour carte à puces

– les divers systèmes de porte-monnaie électronique bien que maintenant il y en ait de moins en moins (Proton au Bénélux, Geld-Karte en Allemagne), suppléé par des transactions EMV petit montant ;
– le système de santé SESAM-VITALE en France ;
– les différents systèmes pour la télévision à péage, utilisant la carte comme porte-clé cryptographique. La carte calcule les clés de désembrouillage d'images en déchiffrant des **mots de contrôle** envoyés périodiquement dans la trame d'émission.

5.3 Systèmes d'exploitation ouverts

Une carte à puces à OS ouvert est en quelque sorte **une carte programmable**, c'est-à-dire une carte dans laquelle il est possible, en plus d'exécuter du code natif suivant le même principe que celui des systèmes fermés, soit de télécharger du code natif après délivrance de la carte, soit de programmer un interpréteur de commande particulier, et donc un jeu d'instructions complémentaire et des mécanismes de gestion de fichiers et/ou programmes additionnels. Cette définition est essentiellement théorique et se heurte immédiatement aux contraintes inhérentes à la sécurité et à l'espace mémoire disponible (*i.e.* la taille de la mémoire programmable, en l'occurrence la taille de la flash le plus souvent) dès que l'on veut la mettre en pratique.

5.3.1 Téléchargement de code natif

Le téléchargement de code natif ne peut être mis en œuvre de manière opérationnelle que si le composant de la carte dispose de **tous les mécanismes de protection mémoire et de supervision adaptés**. Dans la négative, il est très difficile de restreindre *a priori* les

fonctions d'accès aux mémoires du code embarqué. Lorsque le composant offre les garanties adaptées, il y a deux méthodes utilisables :

– télécharger tout le programme dans un fichier élémentaire EF, dont la structure est exécutable. Les paramètres d'exécution du code sont alors passés à la carte *via* la commande EXECUTE, éventuellement précédée d'étapes d'authentification ;
– utiliser l'option commande spécifique ou ASC de la norme NF EN 726-3 (Systèmes de cartes d'identification – Cartes à circuit intégré et terminaux pour les télécommunications – Partie 3 : applications de la carte indépendante des applications) et loger toute l'application (incluant tous ses fichiers et commandes propriétaires) dans un unique fichier DF, dont l'espace mémoire est contrôlé par l'OS. La sélection du fichier en question et l'envoi à la carte d'une commande téléchargée (reconnue alors comme telle par la carte) permettent alors de déclencher l'exécution du code téléchargé.

En fait, ces mécanismes sont le plus souvent utilisés par les industriels de la carte eux-mêmes (en environnement sécurisé, par exemple) pour mettre à jour, voire *patcher* le code embarqué de la carte. Le terme *patch* signifie : modifier le comportement de la carte à puces sans changer le binaire originel (l'OS). Cette technique permet par exemple de limiter le coût d'une re-certification. Seul le patch a besoin d'être évalué le reste du binaire étant inchangé.

Le *patch* était très utilisé lorsque les OS étaient essentiellement en ROM. C'était la seule façon de modifier le comportement d'un OS sachant que l'OS était en zone immuable. Avec les nouvelles technologies et la flash, le « patch » reste encore très utilisé pour éviter de prendre un risque sur une recompilation de tout le binaire, et donc limiter les modifications.

5.3.2 Systèmes d'exploitation à interpréteurs

Le concept d'interpréteur et de machine virtuelle n'est pas nouveau en informatique. La première implémentation d'un tel environnement embarqué dans une carte date de 1996 : le résultat des travaux exploratoires d'une équipe de Schlumberger.

Les principaux objectifs supportant les systèmes d'exploitation à interpréteurs sont les suivants :

– permettre l'écriture d'un code unique, portable immédiatement sur tout nouvel environnement hardware (concept de *Write Once, Run Everywhere*) ;
– permettre l'écriture de code applicatif téléchargeable par des tierces parties n'ayant pas nécessairement une connaissance approfondie de la technologie des cartes ;
– garantir un haut niveau de sécurité, assurant une isolation parfaite entre applications.

À ce jour, il ne reste que 2 principaux systèmes ouverts à interpréteur qui ont été développés pour la carte à puce :

– le plus important est le système JavaCard développé et soutenu par un consortium de plus de 40 industriels réunis au sein du Javacard forum ;
– l'autre est le système Multos, développé initialement par le consortium MAOSCO.

Conceptuellement, ces systèmes sont très proches. Seul JavaCard a connu un réel succès commercial, en raison notamment de son acceptation par l'ETSI comme support au développement de la carte SIM, puis USIM, VISA l'a également adopté très rapidement pour les cartes bancaires avant que tous les autres ne suivent.

Nous nous limiterons dans la suite de ce paragraphe à **JavaCard**, comme élément représentatif des systèmes à interpréteur.

Le système JavaCard (et, par extension, les OS pour cartes à interpréteur JavaCard) est organisé autour de trois composants fondamentaux :

– le langage Javacard,
– la chaîne de développement d'applications autour du langage JavaCard,
– l'environnement embarqué d'exécution ou JCRE (*JavaCard Run Time Environment*).

Le langage JavaCard

Le langage JavaCard est un sous-ensemble du langage Java (paquetage *Java.lang*), il ne supporte donc qu'un nombre restreint d'éléments du langage Java. On peut citer en exemple toutes les fonctions graphiques et le chargement dynamique de classes.

En raison de la nature des cartes à puces, le langage JavaCard supporte deux types d'objets, des objets de type persistant (*persistent objects*), stockés dans la mémoire non volatile (EEPROM/FLASH), et des objets volatiles (*transient objects*), logés dans la mémoire RAM, et dont l'image disparaît à chaque utilisation de la carte.

La chaîne de développement JavaCard

Compte tenu de la puissance limitée de traitement d'une carte à puces, la machine virtuelle JavaCard (JVM) est constituée par deux entités distinctes :

– l'une, logée sur une station de travail ou un ordinateur personnel (***off-card JVM***) ;
– l'autre, embarquée dans la carte (***on-card JVM***).

Une application JavaCard est d'abord développée comme un ensemble de fichiers Java, puis compilée par un compilateur Java standard qui produit des fichiers .class usuels. Un programme additionnel, dénommé ***converter***, implémente une partie de la machine virtuelle, vérifie la compatibilité des classes avec le langage JavaCard, charge et lie tous ces objets dans un paquetage. Le résultat de ces opérations est stocké dans un fichier standardisé appelé. ***CAP*** (***Converted Applet***). Un fichier dit d'exportation (.exp), contient les déclarations des différents éléments du paquetage afin d'en permettre une utilisation ultérieure par un autre paquetage, lors de l'opération de conversion.

Le fichier ***.CAP*** est chargé dans la carte grâce à la collaboration de deux entités : le ***off-card installation program*** et le ***on-card installer***. Cette opération consiste en la segmentation du fichier en une série d'APDUs qui sont signés (et parfois chiffrés) à l'aide de clés secrètes afin d'authentifier leur origine.

Un interpréteur Java (***interpreter***) logé dans la carte réalise l'exécution du *byte code* lors de l'activation d'un applet. Une ***applet*** est un ensemble de classes regroupées dans un même fichier *.CAP*. Elle est identifiée de manière unique par un AID (*Application Identifier*, un nombre de 16 octets) qui, conformément à la norme ISO 7816-5 (voir annexe de la partie documentation de l'article), comporte un préfixe de 5 octets (RID – *Resource IDentifier*) et une extension propriétaire de 11 octets (PIX – *Proprietary Identifier eXtension*). Une application JavaCard est alors articulée autour de l'héritage (au sens Java du terme) d'une applet.

Le JCRE

Le *JavaCard Run Time Environment* (JCRE) est un ensemble de composants résidant dans la carte. Le JCRE est responsable de la gestion des ressources de la carte, de la communication réseau, de l'exécution des applets, du système de la carte et de la sécurité des applets. Il comprend à son tour les éléments suivants :

1) ***Installer***, un bloc réalisant le chargement d'une applet dans la carte ;

2) ***APIs***, un ensemble de quatre paquetages nécessaires à l'exécution d'une applet :

– java.lang (langage),
– javacard.framework (environnement),
– javacard.security (sécurité),
– javacardx.crypto (cryptographie) ;

3) des composants spécifiques à une application particulière ;

4) un ensemble de classes (***system classes***) réalisant les services de base, tels que :

– la gestion des applets,
– la gestion des transactions,
– les communications,
– autres…

5) la ***machine virtuelle*** et les méthodes natives associées.

Le dialogue avec l'entité JCRE est assuré à l'aide d'APDUs : une classe dite ***APDU*** comporte toutes les méthodes nécessaires à la réception des commandes et à la génération des réponses. Comme pour tous les OS pour cartes à puces, la norme JavaCard introduit une notion d'opération atomique relativement à la mémoire non volatile.

6. Normalisation

6.1 Information globale et situation

Le but de la normalisation est que, pour une application donnée, toute carte fonctionne de manière identique sur tout terminal qui lui est associé. Les normes spécifient donc l'interface entre le microcircuit et le monde extérieur : caractéristiques physiques, caractéristiques électriques, remise à zéro des circuits, protocoles d'échanges, définition et codage des commandes, déroulement d'une transaction ou d'une session. Les normes de base sont dites inter-industrielles, car elles sont générales et indépendantes des applications.

C'est dès 1980 que les travaux normatifs ont démarré en France au sein de l'AFNOR et, en octobre 1981, le Comité Technique TC97 de l'ISO chargé des systèmes d'informations adoptait un nouveau sujet de travail : normaliser l'interface des cartes à microcircuits à contacts.

6.1.1 Fondement normatif

La normalisation internationale de la carte à puces se poursuit depuis cette date dans le cadre du JTC1/SC17/WG4 avec comme principaux acteurs : France, Japon, Allemagne, États-Unis, Norvège et Royaume-Uni (plus d'une trentaine de pays au total). Les membres de « liaisons » tels que Intamic (*International Association for Microcircuit Cards*), Mastercard, Visa sont également très actifs. Jusqu'à ce jour, le WG4 a déjà réalisé en particulier 13 standards internationaux génériques rassemblés au sein de la norme ISO IEC 7816, et qui constituent le fondement normatif incontournable de la carte à puces :

– 7816-1 : Partie 1 : *Cartes à contacts – Caractéristiques physiques des cartes* ;

– 7816-2 : Partie 2 : *Cartes à contacts – Dimensions et positions des contacts électriques* ;

– 7816-3 : Partie 3 : *Cartes à contacts – Interfaces électriques et protocoles de transmission* ;

– 7816-4 : Partie 4 – *Organisation, sécurité et commandes pour l'échange* ;

– 7816-5 : Partie 5 : *Enregistrement des fournisseurs d'application* ;

– 7816-6 : Partie 6 : *Éléments de données pour l'échange* ;

– 7816-7 : Partie 7 : *Commandes intersectorielles pour langage de requête de carte structurée* (SCQL) ;

– 7816-8 : Partie 8 : *Commandes et mécanismes pour les opérations de sécurité* ;

– 7816-9 : Partie 9 : *Commandes pour la gestion de cartes* ;

– 7816-10 : Partie 10 : *Signaux électroniques et réponse à réinitialiser pour les cartes synchrones* ;

7816-11 : Partie 11 : *Vérification personnelle par méthodes biométriques* ;

– 7816-12 : Partie 12 : *Cartes à contact de fonctionnement – Interface électrique USB et procédures* ;

– 7816-13 : Partie 13 : *Commandes pour la gestion d'applications dans un environnement multi-applications – Syntaxe et format des commandes et mécanismes cryptographiques*.

Dans le domaine des cartes sans contact, ce même sous-comité SC17 du groupe commun ISO/IEC a publié une norme générique ISO-IEC 14443 *Cartes et dispositifs de sécurité pour l'identification personnelle – Objets sans contact de proximité* en quatre parties pour les **cartes sans contact de proximité** (distance carte lecteur de l'ordre de 10 cm) :

– Partie 1 : *Caractéristiques physiques*,

– Partie 2 : *Interface radio fréquence et des signaux de communication*,

– Partie 3 : *Initialisation et anticollision*,

– Partie 4 : *Protocole de transmission*.

6.1.2 Applications

Au niveau des applications, il est difficile d'être exhaustif, tant l'explosion des marchés a conduit à une prolifération des comités de normalisation *ad hoc* ou officiels ; signalons cependant pour les principaux domaines de la cartes à puces les principaux standards opérationnels mis en œuvre.

Téléphonie mobile

Les standards clefs gouvernant le rôle de la carte SIM dans les réseaux GSM ont été définis à l'ETSI dans de nombreux documents. Parmi ces derniers, les cinq textes suivants forment, à l'origine, une référence de base pour toute implémentation :

– GSM 11-11 : *Spécification de l'interface SIM-ME* ;

– GSM 11-14 : *Spécification SIM Application Toolkit pour l'interface SIM-ME* ;

– GSM 03.19 : *API JavaCard™ de programmation pour la carte SIM Phase 2* ;

– GSM 03.40 : *Réalisation de la fonction Short Message Service (SMS) ; mode Point to Point (PP)* ;

– GSM 03.48 : *Mécanismes de sécurité pour la carte SIM application toolkit Phase 2*.

Ces standards ont été repris ou étendus dans le cadre des groupes de travail constitués par l'ITU pour la définition des systèmes de téléphonie mobile de 3^{e} génération, notamment dans les documents répertoriés dans les familles TS 23.0x, TS.31.1xx et TS 102.2xx pour les systèmes W-CDMA (une vingtaine de documents au total).

Pour la 4G, citons les standards LTE (*Long Term Evolution*) et LTE Advanced.

Domaine bancaire

Les réseaux VISA et MASTERCARD ont établi, avec leurs membres, des spécifications communes, baptisées EMV, permettant l'interopérabilité internationale des transactions de paiement et de retrait par cartes à puces de type interbancaire dans un contexte multi-applicatif. Les normes EMV visent à la fois à permettre le développement de nouveaux services et à améliorer la sécurité des transactions.

La première version du standard EMV a été publiée en 1996. La dernière version, dite EMV 2000, comporte quatre parties :

– Partie 1 : *Caractéristiques mécaniques, électriques, interface logique et protocoles de transport des cartes, commandes et mécanismes de base pour la sélection d'applications* ;

– Partie 2 : *Sécurité et gestion des clefs cryptographiques* ;

– Partie 3 : *Spécification d'applications (en particulier pour les fonctions de débit-crédit)* ;

– Partie 4 : *Exigences au niveau des terminaux supportant les cartes EMV*.

Les banques et la grande majorité des commerçants français suivent les normes EMV. Le déploiement du standard EMV est par ailleurs complet dans les pays européens, asiatiques et du continent sud-américain et en cours dans tous les grands pays, y compris l'Inde et les États-Unis.

Signature électronique

En vue des applications liées au développement de commerce électronique, la Commission Européenne a publié en 1999 une directive pour la signature électronique (1999-93-EC aussi connue sous l'acronyme CEN/ISSS). Cette directive a été transposée dans les états membres, et a donné naissance à toute une famille de normes liées aux fonctions d'identification, authentification et signature digitale dans les cartes à puces, notamment au sein des groupes CEN-TC 224 WG15 pour la définition de la future « carte d'identité du citoyen européen ». L'ETSI a également des actions de standardisation pour les formats et infrastructures de signature électronique pour les applications liées au commerce électronique (ETSI TS 101 733 : *Electronic Signatures and Infrastructures* (*ESI*) ; *CMS Advanced Electronic Signatures* (*CadES*)).

6.2 Caractéristiques physiques des cartes et position des contacts électriques

La norme ISO 7816-1 définit les dimensions des cartes ainsi que leur environnement standard. Dès 1982, l'AFNOR proposait une position et une affectation des contacts. Proche du coin supérieur gauche de la carte, cette position « haute » correspondait à une bonne fiabilité du module électronique. Néanmoins, l'ISO a adopté dans la norme 7816-2 une position plus centrale réalisant le compromis demandé par les Japonais qui possèdent une piste magnétique au même niveau que la position haute.

6.3 Interface électrique des cartes

La norme traite uniquement les cas des cartes à microprocesseur, en laissant de côté les cartes à logique câblée qui possèdent des interfaces de transmission synchrone très particulières, incompatibles entre elles.

C'est la solution française qui a été retenue intégralement dans la norme, ce succès ayant été masqué par les tergiversations sur la position des contacts. Les signaux électriques sont définis statiquement et dynamiquement. Comme la carte est essentiellement amovible, le standard décrit la procédure d'activation et de désactivation des contacts.

Il faut noter que seules les plages de contacts sont définies dans la norme, comme étant des rectangles métalliques de 2 mm x 1,7 mm. La forme du support de ces contacts est laissée libre, à condition que l'on évite les courts- circuits, notamment entre les positions haute et basse.

Une grande bataille eut lieu sur **la fréquence de référence**, qui est la fréquence fournie sur le contact d'horloge (CLK) pour que les données soient échangées à 9 600 bits par seconde sur le contact d'entrée-sortie (E-S ou I-O dans la littérature anglo-saxonne). Il y a lieu de ne pas confondre cette fréquence avec celle qui peut réellement être appliquée sur CLK et qui peut dépasser 20 MHz (voire approcher les 50 MHz sur les derniers processeurs apparus sur le marché).

La fréquence de référence choisie est 3,579 545 MHz, car c'est la plus haute fréquence normalisée inférieure à 4 MHz, fréquence considérée comme limite pour les micro-calculateurs dans les années 80. Il faut insister sur le fait que plus la fréquence de référence est basse, plus rapide est la vitesse relative du microprocesseur.

Pour les valeurs des signaux électriques et leur séquencement, on ne saurait trop insister sur le respect impératif de la norme qui assure non seulement l'interopérabilité, mais aussi la fiabilité de la liaison avec les cartes. Le non-respect de cette norme a conduit à des déboires dans le monde bancaire (en particulier lors du déploiement en France de la carte bancaire à puces) où la remise à niveau des terminaux a demandé plusieurs années !

Dès la mise sous tension, les premières données échangées proviennent de la carte et constituent la **réponse à la remise à zéro** (RAZ) ou ATR (*Answer to Reset*). Ce message initial, qui doit être envoyé par toute carte, décrit le mode de transmission des caractères, la fréquence d'horloge, la vitesse de transmission et le type de protocole supporté. Cette séquence comporte aussi des informations propres au microcircuit contenues dans une suite d'octets historiques.

6.4 Protocoles d'échanges

6.4.1 Protocole d'échanges par caractères (*T* = 0)

Ce protocole, développé à l'origine entre Bull et le CCETT, puis normalisé au plan international en 1989, a été conçu pour minimiser le prix des cartes, tout en présentant de bonnes performances, aussi est-il largement employé. Il utilise une transmission asynchrone par caractères. Cela signifie simplement que les erreurs de transmission sont gérées au niveau de chaque caractère afin d'être corrigées immédiatement par répétition, ce qui réduit considérablement la taille de la mémoire tampon.

La trame générale utilisée comprend un en-tête de 5 octets : classe (CLA), instruction (INS), paramètres (P1, P2, P3).

Généralement, P1 et P2 donnent une référence, P3 indique la longueur du champ de données attendu. Les données peuvent être **entrantes** ou **sortantes**, en fonction de l'instruction considérée. Le dernier champ contient deux octets d'état ME1, ME2 qui donnent un compte-rendu d'exécution.

Le mécanisme de reprise d'erreur permet de répéter un caractère erroné immédiatement après sa détection. Lorsque le récepteur détecte une erreur, il force la ligne pour signaler à l'émetteur qu'il faut répéter immédiatement le caractère transmis.

Un service important fourni par le protocole *T* = 0 est l'asservissement de vitesse au niveau caractère, qui peut s'avérer très utile pour réguler le flux des données transmises en fonction de la vitesse de traitement, et en utilisant une RAM de taille très réduite, donc une carte peu coûteuse.

6.4.2 Protocole d'échanges par blocs (*T* = 1)

Sous la poussée des Japonais et des Allemands un protocole par blocs a été inclus dans la norme. Quel est l'intérêt de cet autre protocole ?

Les défenseurs du protocole avancent l'argument suivant : un protocole d'échange académique doit permettre à un ordinateur distant de s'adresser directement à une carte particulière, sans faire appel à une intelligence locale. Ainsi, une protection contre les erreurs de transmission doit se faire au niveau d'un bloc entier de caractères. Le principe d'un protocole bloc étant adopté, d'autres services viennent s'ajouter aux précédents de façon à pouvoir pleinement utiliser cette structure :

– détection des erreurs au niveau bloc par deux octets de redondance ;
– adressage de plusieurs canaux à l'intérieur d'une même carte ;
– échange bidirectionnel de caractères dans la même commande ;
– allocation dynamique de la longueur maximale des blocs ;
– chaînage des commandes dans plusieurs blocs.

Un bloc est formé de trois parties :

– le prologue est constitué d'un en-tête obligatoire de 3 octets et il comprend l'adresse NAD, le caractère de contrôle PCB et la longueur du champ de données LEN ;
– les informations constituent un champ de données optionnel ;
– l'épilogue obligatoire peut contenir 1 octet (LRC – *Linear Redomption Check*) ou 2 octets (CRC *Cyclic Redomption Check*).

L'octet PCB est utilisé pour gérer le transport du bloc, il permet de distinguer et de numéroter différents types de blocs. Les données peuvent théoriquement avoir une longueur comprise entre 0 et 254.

La méthode d'acquisition du bloc n'est pas précisée dans la norme, et aucune erreur n'est donc corrigée au niveau caractère. Une erreur de parité caractère ou de *checksum* ne peut être traitée qu'après l'arrivée du bloc, et provoque l'envoi d'un acquittement négatif, que suivra la répétition du bloc entier.

Il existe 3 types de blocs :

– les blocs I numérotés sur un bit qui ont en général un champ d'informations, ils sont utilisés par la couche application et acquittent positivement le bloc précédent ;
– les blocs R qui acquittent les blocs et contiennent le numéro du bloc I attendu ;
– les blocs S de supervision qui sont utilisés pour échanger des données de service.

La totalité d'un bloc doit être mise dans une mémoire tampon avant traitement. Lorsque les informations à échanger dans une commande sont trop longues pour le tampon courant, il faut les répartir dans plusieurs blocs chaînés entre eux.

Comme dans tout protocole, le fonctionnement normal sans erreur ne présente aucune difficulté particulière. La complexité du protocole se manifeste dès lors que l'on considère la détection et la reprise des erreurs de transmission.

6.5 Jeu de commandes inter-industries

6.5.1 Architecture générale et sécurité

Il est très vite apparu aux normalisateurs qu'un jeu de commandes ne peut être défini que si l'on connaît les structures ou les objets sur lesquels agissent ces commandes. Lorsque l'on conçoit un système d'exploitation, il est classique d'opérer sur une architecture arborescente hiérarchisée de fichiers, comme c'est le cas pour Windows ou Linux, mais pour la carte à microcalculateur les impératifs sécuritaires doivent être pris en compte *a priori* dans la structure et les méthodes d'accès. C'est pourquoi la norme ISO 7816 introduit l'architecture de sécurité des cartes et la messagerie correspondante.

Organisation des fichiers

La racine est constituée par le fichier maître MF, créé à la personnalisation de la carte. Le MF peut contenir des fichiers élémentaires EF ou des fichiers dédiés DF pouvant eux-mêmes contenir des EF. Les fichiers peuvent être référencés par un identificateur, par un chemin ou par un nom. Les fichiers élémentaires peuvent avoir une structure de type **transparent** (ne contenant des données binaires), **linéaire** (contenant une structure d'enregistrements de longueur variable ou fixe), **cyclique** (le rangement des enregistrements pouvant être organisé séquentiellement jusqu'au nombre maximal d'enregistrements déclaré, après quoi, l'ajout de nouveaux enregistrements s'effectue par ré-écriture à partir du premier enregistrement).

Les **attributs de sécurité** associés à chaque fichier et introduits lors de leur création fixent les conditions d'accès permettant d'exécuter des commandes sur ces fichiers. La gestion de la sécurité est fondée sur les principes suivants :

– un accès licite augmente les droits d'accès ;
– une opération n'est autorisée que lorsque les droits acquis sont égaux ou supérieurs aux attributs de sécurité ;
– l'héritage des droits se transmet de parent à enfant mais non l'inverse.

Les **mécanismes généraux de sécurité** concernent :

– l'authentification interne ou externe des cartes ;
– la présentation des mots de passe qui protègent les porteurs ;
– le chiffrement et le déchiffrement des données par la carte.

Les **statuts de sécurité** représentent l'état courant dans lequel se trouve la carte après une opération. Ces statuts sont au niveau MF, DF et pour chaque commande exécutée.

La **messagerie de sécurité** assure l'authentification ou la confidentialité à l'aide de mécanismes cryptographiques. Les sommes de contrôle et les signatures numériques sont supportées par le format et les éléments de données.

Structure des commandes

En principe, le protocole applicatif est transparent au protocole de transport, conformément au modèle OSI. C'est pourquoi la norme 7816 définit des unités de données de protocole applicatif (APDU), qui contiennent des paires commande-réponse (CRP). Il y a quatre structures de commandes illustrées par la figure **17**. C'est la connaissance de la longueur totale de l'APDU qui permet au récepteur de déterminer le type de CRP dont il s'agit.

Structure des réponses et statuts d'état

La structure des réponses est unique, elle comprend un champ de données et un champ contenant deux octets d'état SW1 et SW2. D'une façon générale, les valeurs de SW1 et SW2 sont définies au titre de la syntaxe de chaque commande.

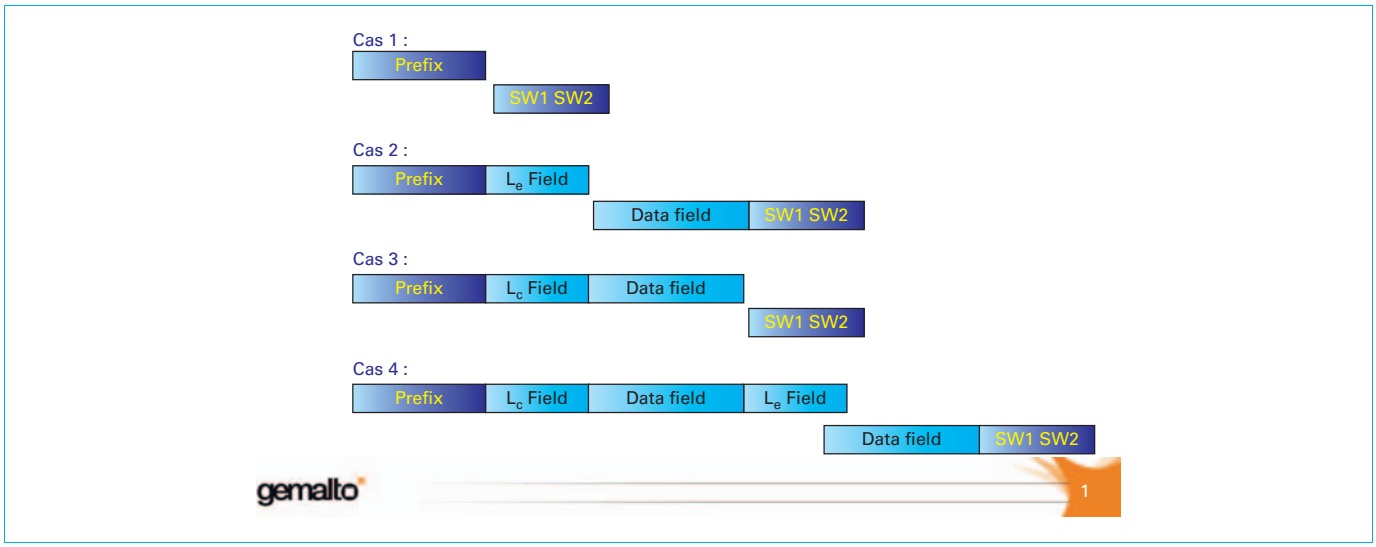


Figure 17 – Les quatre structures de commande

Structure des champs de données

Bien que la norme accepte des champs de données quelconques propriétaires, elle définit trois méthodes de présentation des données fondées sur une structure TLV (étiquette, longueur, valeur). On distingue les objets ASN.1 et les objets TLV simples ou compacts.

Les objets ASN.1 proviennent des normes ISO 8824 (Systèmes de traitement de l'information – Interconnexion de systèmes ouverts – Spécification de la notation de syntaxe abstraite numéro 1) et 8825 (Systèmes de traitement de l'information – Interconnexion de systèmes Ouverts – Spécification des règles de codage de base pour la notation de syntaxe abstraite numéro 1). Il existe des étiquettes contextuelles correspondant à des objets qui dépendent de la façon dont on les a obtenus et des étiquettes applicatives pour des objets indépendants du contexte. L'étiquette détermine aussi si l'objet est « construit » à l'aide de plusieurs objets ASN.1 consécutifs ou si l'objet est « primitif » en transportant des objets TLV simples.

6.5.2 Jeu de commandes

La liste des commandes définies en 1994 peut regrouper en quatre familles :

– la **méthode d'accès** constituée par les commandes **Select File**, **Read**, **Write**, **Update** et **Erase** qui peuvent s'appliquer aussi bien aux fichiers à structure transparente qu'à ceux organisés en enregistrements. La commande **Log** (dénommée aussi ***Append***) place un enregistrement après le dernier mémorisé ;
– les **commandes de sécurité** permettent de présenter un mot de passe (**Verify**), d'obtenir un nombe aléatoire (**Get Challenge**), ou encore d'authentifier une carte ou un terminal (**Authenticate**) ;
– les **commandes orientées « transmission » Get Response** et **Envelope** permettent de transporter des APDU partiels ou complets sortant de la carte ou allant vers cette dernière. Le **Get Response** est surtout utilisé pour récupérer les réponses en protocole $T = 0$. La gestion des canaux est assurée par la commande **Manage Channel**. Elle permet d'ouvrir et de fermer un canal ;
– les **commandes génériques Put** et **Get** écrivent ou renvoient respectivement des données non référencées. Elles sont destinées à cacher au monde extérieur la structure des fichiers, elles ne peuvent donc avoir qu'un comportement propriétaire assez antinomique avec une norme.

7. Perspectives d'avenir

Des progrès fantastiques ont été accomplis dans le monde des semi-conducteurs depuis les années 1970, le doublement de la puissance des processeurs ayant lieu tous les deux ans. Aujourd'hui, le perfectionnement des techniques de gravure permet de descendre à des finesses de gravure de 10, 7 voire même 5 nm. Le 7 nm est entré en production en 2018. Cette évolution appelée Loi de Moore s'est poursuivie jusqu'à aujourd'hui avant de ralentir et de céder probablement à de nouvelles approches mettant en œuvre des supraconducteurs, photons, ions et structures 3D.

Le marché de la carte à puces s'est lui aussi approprié cette révolution technologique et déploie massivement des puces avec des gravures de 90 nm, voire 45 nm.

La taille des puces pour cartes limitée à 25 mm^2, principalement en raison des contraintes physiques liées à l'encartage, n'est donc plus un facteur critique puisque plusieurs dizaines voire centaines de millions de transistors pourront y être intégrés, amenant la puissance bien au-delà de celle des premiers PC !

Ces progrès d'intégration permettent ainsi à l'industrie de la carte à puces de développer un modèle économique très compétitif, et d'améliorer les performances d'un même produit de 15 à 20 % par an. En parallèle, ces technologies fortement submicroniques ou nanométriques apportent une diminution des énergies et des tensions d'alimentation. Les cartes pourront fonctionner autour de 1 V, voire moins, ce qui favorisera la prolifération des cartes sans contact, en particulier pour les applications liées au transport, à la banque, à l'identité ou à l'identification, et à la téléphonie mobile (comme en témoignent les applications actuelles basées sur la technologie NFC ou *Near Field Communication*).

Les densités des mémoires non volatiles ont progressé de façon significative, permettant d'augmenter les capacités des cartes à puces bien au-delà de 1 Mbit. Une évolution majeure de ces mémoires non volatiles, le cœur même des cartes à puces, a déjà eu lieu avec l'arrivée en production de masse des mémoires flash. Ces nouvelles générations de mémoires non volatiles ont un temps d'accès comparable à celui des SDRAM et un encombrement équivalent aux mémoires ROM. Des mégaoctets, voire des gigaoctets, pourront ainsi être stockés quasiment sans aucune consommation d'énergie.

Une deuxième révolution concerne l'évolution des protocoles de communication supportés par la carte : illustrés par le passé avec l'apparition du protocole USB 1.0 supporté en parallèle aux protocoles ISO, ils sont beaucoup plus rapides et plus efficaces que ceux-ci. Le changement de protocole de transmission entre la carte et le monde extérieur permet d'augmenter la vitesse de transfert pour atteindre des vitesses de l'ordre de plusieurs mégabits par seconde. Cela est réalisé par l'intégration dans la carte de protocoles tels que SWP (*Single Wire Protocol*) ou SPI. Les deux protocoles sont rapides, jusqu'à 20 $Mbits.s^{-1}$ pour SPI et 1,5 $Mbit.s^{-1}$ pour le SWP, et leur efficacité est supérieure à 80 %.

Les débats liés aux architectures de composants pour les cartes à puces (architecture 32 bits versus 8 bits, caches intégrés, pipeline…) sont relégués aux oubliettes de l'histoire.

Ces progrès permettent donc de prévoir plusieurs types d'évolutions pour les cartes à puces et pour les briques technologiques associées :

- tout d'abord l'accélération du développement d'applications de masse autour des cartes de paiement, des passeports, des cartes d'accès réseau, des cartes d'identité et des permis de conduire électroniques. Le coût de la puce n'est plus un facteur de limitation et apporte la sécurité indispensable pour l'identification, l'authentification et la validation des transactions ;
- la deuxième évolution exploitera les possibilités d'intégration réseau en augmentant les puissances de traitement et les capacités de mémoire. Les langages évolués et les bibliothèques associées se généraliseront pour la couche applicative des cartes, offrant ainsi aux programmeurs d'applications la visibilité d'une plate-forme ouverte, permettant d'en démultiplier les usages. En ajoutant le formidable bond de puissance réalisé par le microprocesseur (CPU, mémoire, protocoles de communication…) et le logiciel embarqué, la carte est devenue peu à peu le support de la mobilité personnelle. Elle pourra contenir l'ensemble des paramètres d'accès au réseau. Supportant directement les protocoles de l'Internet (TCP/IP, FTP…) et les modes client-serveurs associés, la carte du futur deviendra de plus en plus intelligente et offrira un très fort taux d'interactivité. Elle servira de cache pour les transactions au niveau utilisateur. Elle disposera naturellement d'un très haut niveau de sécurité. Tous les mots de passe de l'utilisateur seront conservés dans la carte, une copie de celle-ci étant conservée dans un endroit protégé. Dans certains schémas d'architecture (par exemple, basés sur des mécanismes de fédération d'identité de type *Liberty Alliance*), la carte est déjà le réceptacle de tous les attributs d'accès ou de profils utilisateur de son porteur, et elle est capable de réaliser elle-même, de manière automatisée et complètement sécuritaire, l'accès aux services souscrits, tout en offrant à l'utilisateur des moyens extrêmement puissants de contrôler les données relatives à sa vie privée ;
- une troisième évolution prévisible se situe dans le bas de la gamme que les microprocesseurs ne pouvaient atteindre jusqu'à présent. C'est probablement la plus spectaculaire et la plus prometteuse des tendances, car les besoins accrus de fonctionnalités et de sécurité condamnent à terme les puces en logique câblée, tandis que les microprocesseurs seront capables de les remplacer pour un prix équivalent. Cette

dernière tendance ouvre le marché gigantesque des cartes privatives, des transports en commun, la protection de l'IoT (l'internet des objets), ainsi que le domaine des dispositifs d'identification électroniques de type RFID ;

- une dernière évolution vient de l'exploitation des technologies des cartes à puces au sein même de réalisation plus complexes. Il s'agira de garantir à l'intérieur d'un processeur plus puissant, une zone où les secrets et les applications sont parfaitement protégés et exécutés. Pour les métiers de la carte à puces, ce marché reste à conquérir.

On voit que les potentialités des cartes à puces sont très importantes et qu'elles continueront de jouer un rôle de premier plan pour assurer la protection des informations dans nos sociétés futures.

Sous ses divers facteurs de forme, comme par exemple un élément sécurisé discret, les cartes à puces sauront aussi être un allié indispensable des téléphones portables pour pouvoir atteindre des niveaux de sécurité très importants.

8. Termes ou abréviations

Abréviation	Définition	Traduction
ACL	*Access Control List*	Liste de contrôle d'accès
AES	*Advanced Encryption Standard*	Standard d'encryptage avancé
APDU	*Application Protocol Data Unit*	Unités de données de protocole applicatif
ASN	*Abstract Syntax Notation one*	–
ATR	*Answer To Reset*	Réponse à l'initialisation
CBC	*Cipher Bloc Chaining*	–
CCETT	Centre commun d'études de télévision et télécommunications	–
CMOS	*Complementary Metal Oxyde Semiconductor*	–
CRP	–	Paire de commande-réponse
DES	*Data Encryption Standard*	Standard d'encryptage de données
DF	–	Fichier dédié
DPA	*Differential Power Analysis*	Analyse par différence de consommation
DRAM	*Dynamic Random Access Memory*	Mémoire RAM dynamique
DSP	*Digital Signal Processor*	Processeur de signaux numériques
ECB	*Electronic Code Book*	–
EEPROM	*Electrically Erasable Programmable ROM*	–
EF	–	Fichier élémentaire
ETSI	*European Telecommunications Standards Institute*	–
EPROM	*Erasable Programmable ROM*	–
EMV	*Europay Mastercard Visa*	–
GSM	*Global System for Mobile communications*	Groupe spécial systèmes mobiles
IGC	–	Infrastructure de Gestion de Clés
ITU	–	Union internationale des télécommunications
JRCE	*JavaCard Run Time Environment*	–
MF	–	Fichier maître
MMU	*Memory Management Unit*	Unité de gestion de la mémoire
MOS	*Metal Oxyde Semiconductor*	–

Abréviation	Définition	Traduction
MRAM	*Magnetoresistive RAM*	–
NVM	*Non Volatile Memory*	Mémoire non volatile
OS	*Operating System*	Système d'exploitation
PC-RAM	*Phase-Change RAM*	–
PKI	*Public key infrastructure*	–
PROM	*Programmable ROM*	Mémoire morte programmable
RAM	*Random Access Memory*	Mémoire à accès aléatoire
RAS	–	Rivest, Shamir, Adleman (nom d'un algorithme)
RAZ	–	Remise à zéro
RISC	*Reduced Instruction Set Computer*	–
ROM	*Read Only Memory*	Mémoire morte
SE	*Secure Element*	Elément de sécurité
SIM	*Subscriber Identity Module*	–
SPOM	*Self Programmable One chip Microcomputer*	Microcalculateur autoprogrammable monolithique
SRAM	*Static Random Access Memory*	Mémoire RAM statique
TAB	*Tape Automated Bonding*	–
TLV	–	Étiquette, longueur, valeur (structure)
USIM	*Universal Subscriber Identity Module*	–
IoT	*Internet of Object*	Internet des objets

Cartes à puces

Technologie et cybersécurité

par **Jean-Pierre TUAL**
Ancien directeur des relations industrielles,
Direction technologie et innovation,
Gemalto
Auteur de la version originale de l'article 2007

Stéphane GRELLIER
Mobile software security & services manager,
Gemalto, Meudon, France
Auteur de la version actualisée de 2019

Joseph LEIBENGUTH
Physical document security R&D product director – Technical advisor,
Gemalto,
Saint-Cloud, France
Auteur de la version actualisée de 2019

et **Philippe PROUST**
Embedded & core security director,
Gemalto, Géménos , France
Auteur de la version actualisée de 2019

Sources bibliographiques

[1] RIVEST (R.L.), SHAMIR (A.) et ADLEMAN-COMMUN (L.). – *A method for obtaining digital signatures and public-key crypto systems*. ACM, vol. 21, n° 2, p. 120-126 (1978).

[2] UGON (M.) et GUILLOU (L.C.). – *Les cartes à puces*. La Recherche n° 176 (1986).

[3] RANKL (W.) et EFFING (W.). – Smart card Handbook John Wiley & Sons (2002).

[4] GUILLOU (L.C.) et QUISQUATER (J.J.). – *A practical Zero Knowledge protocol fitted to security microprocessor minimising both transmission and memory*. Proc. Eurocrypt. Springer Verlag (1988).

[5] GUEZ (F.), ROBERT (C.) et LAURET (A.). – *Les cartes à microcircuit*. Masson (1988).

[6] *Smart Card 2000*. Édité par D. Chaum-North Holland (1989) et (1991).

[7] GUILLOU (L.C.), UGON (M.) et QUISQUATER (J.J.). – *The smart card. A standardised security device dedicated to cryptology*. Contemporary cryptology, chap. 12. – IEE Press (1992).

À lire également dans nos bases

CHASSÉ (G.). – *Cryptographie – Algorithmes*. [AF 173] (2000).

CHASSÉ (G.). – *Cryptographie – Mathématiques*. [AF 172] (2000).

FOUQUE (P.-A.). – *Cryptographie appliquée*. [H 5 210] (2003).

POUPON (G.). – *Procédés de packaging et d'interconnexion de composants électroniques*. [E 3 401] (2016).

SKOTNICKI (T.). – *Transistor MOS et sa technologie de fabrication*. [E 2 430] (2000).

TECHNIQUES DE L'INGÉNIEUR

GAGNEZ DU TEMPS ET SÉCURISEZ VOS PROJETS EN UTILISANT UNE SOURCE ACTUALISÉE ET FIABLE

Techniques de l'Ingénieur propose la plus importante collection documentaire technique et scientifique en français !

Grâce à vos droits d'accès, retrouvez l'ensemble des **articles et fiches pratiques de votre offre, leurs compléments et mises à jour,** et bénéficiez des **services inclus.**

RÉDIGÉE ET VALIDÉE PAR DES EXPERTS	MISE À JOUR PERMANENTE	100 % COMPATIBLE SUR TOUS SUPPORTS NUMÉRIQUES	SERVICES INCLUS DANS CHAQUE OFFRE

- **+ de 350 000 utilisateurs**
- **+ de 10 000 articles de référence**
- **+ de 80 offres**
- **15 domaines d'expertise**

- **Automatique - Robotique**
- **Biomédical - Pharma**
- **Construction et travaux publics**
- **Électronique - Photonique**
- **Énergies**
- **Environnement - Sécurité**
- **Génie industriel**
- **Ingénierie des transports**
- **Innovation**
- **Matériaux**
- **Mécanique**
- **Mesures - Analyses**
- **Procédés chimie - Bio - Agro**
- **Sciences fondamentales**
- **Technologies de l'information**

Pour des offres toujours plus adaptées à votre métier, découvrez les offres dédiées à votre secteur d'activité

Depuis plus de 70 ans, Techniques de l'Ingénieur est la source d'informations de référence des bureaux d'études, de la R&D et de l'innovation.

www.techniques-ingenieur.fr

CONTACT : Tél. : + 33 (0)1 53 35 20 20 - Fax : +33 (0)1 53 26 79 18 - E-mail : infos.clients@teching.com

TECHNIQUES DE L'INGÉNIEUR

LES AVANTAGES ET SERVICES compris dans les offres Techniques de l'Ingénieur

ACCÈS

Accès illimité aux articles en HTML
Enrichis et mis à jour pendant toute la durée de la souscription

Téléchargement des articles au format PDF
Pour un usage en toute liberté

Consultation sur tous les supports numériques
Des contenus optimisés pour ordinateurs, tablettes et mobiles

SERVICES ET OUTILS PRATIQUES

Questions aux experts*
Les meilleurs experts techniques et scientifiques vous répondent

Articles Découverte
La possibilité de consulter des articles en dehors de votre offre

Dictionnaire technique multilingue
45 000 termes en français, anglais, espagnol et allemand

Archives
Technologies anciennes et versions antérieures des articles

Impression à la demande
Commandez les éditions papier de vos ressources documentaires

Alertes actualisations
Recevez par email toutes les nouveautés de vos ressources documentaires

*Questions aux experts est un service réservé aux entreprises, non proposé dans les offres écoles, universités ou pour tout autre organisme de formation.

ILS NOUS FONT CONFIANCE

AgroParisTech · AIRBUS · cea · CESI Enseignement Supérieur et Formation Professionnelle · cnes Centre National d'Études Spatiales · CSTB le futur en construction · ENGIE · GE · L'ORÉAL · Leroy-Somer · École Polytechnique Université Paris-Saclay · SAFRAN Aerospace Defence Security · Sagemcom · SAINT-GOBAIN · Schneider Electric · THALES · TOTAL · Valeo · VEOLIA · Wallonie · SPW Service public de Wallonie

www.techniques-ingenieur.fr

CONTACT : Tél. : + 33 (0)1 53 35 20 20 - Fax : +33 (0)1 53 26 79 18 - E-mail : infos.clients@teching.com