

Avant-propos

« Pour les aider à se débarrasser de tout ce qu'il ne fallait pas savoir, les Shadoks avaient créé l'Antimémoire. C'était un grand machin à base de mécaniques subtiles telles que poubelles à tiroir, concasseurs de connaissances, broyeurs à savoir, etc. »

Les Shadoks¹,
concernant l'ordinateur

Les systèmes informatiques (matériels et logiciels) deviennent de plus en plus complexes, enfouis et transparents. Il devient alors difficile de retrouver les concepts de base pour comprendre finement leur fonctionnement. Pour y arriver, une voie est de s'intéresser à l'histoire du domaine. Une deuxième voie est de s'imprégner de la technologie par la lecture de brevets et de feuilles de caractéristiques (*datasheets*) de composants électroniques. Une dernière, et non des moindres, est de lire des articles de recherche. Nous avons essayé de suivre ces trois voies tout au long de la rédaction de cette série d'ouvrages, dans le but d'expliquer le fonctionnement matériel de la mémoire, sous sa forme d'un composant ou d'un (sous-)système complet comme un périphérique de mémoire de masse.

1. Créés par Jacques Rouxel, les Shadoks sont les personnages principaux d'un dessin animé expérimental issu du Service de la recherche de l'Office de radiodiffusion-télévision française (ORTF). Ce feuilleton quotidien culte d'une durée de 2 minutes était diffusé sur la première chaîne de l'ORTF (la seule à l'époque !) à partir de 1968. Ces oiseaux étaient dessinés de manière schématique et rapide grâce à l'appareil expérimental appelé *animographe*. Les Shadoks sont ridicules, bêtes et méchants. Ils sont dotés de moyens intellectuels parfaitement anormaux. Par exemple, ils sont connus pour le fait de pomper, on ne sait plus pour quelle raison ! Le vocabulaire de leur langue se résume à quatre mots : GA, BU, ZO et MEU qui sont aussi les quatre chiffres de leur système de numération (base 4) et les notes musicales de leur gamme tétra-tonique. Leur philosophie se résume à des devises célèbres comme celle citée dans cet ouvrage.

À propos de cette série et de cet ouvrage

Dans cette deuxième série d'ouvrages, nous nous intéressons à une fonction essentielle de l'ordinateur qui est la mémorisation, les deux autres fonctions étant le calcul, abordé dans les volumes 1 à 5 de *Le microprocesseur* (Darche 2021a, 2021b, 2021c, 2021d, 2021e), et la communication. L'ordre de présentation des modèles de mémoires a été choisi par rapport à la facilité d'interfaçage ou selon une complexité croissante des concepts de mémorisation.

Ce premier volume est l'objet des prolégomènes à ce domaine. Après une présentation des caractéristiques principales d'une mémoire générique, des technologies passées et présentes et du concept de hiérarchie des mémoires, nous nous intéresserons plus particulièrement aux modèles à semi-conducteurs (*solid-state memory*) à accès aléatoire.

Les volumes 2 et 3 présentent les différentes catégories de mémoires vives à semi-conducteurs à accès aléatoire ou RAM (*Random Access Memory*), c'est-à-dire les modèles statique et dynamique, en version asynchrone, respectivement nommés SRAM (*Static RAM*) et DRAM (*Dynamic RAM*). En particulier, ils détaillent leurs caractéristiques temporelles, électriques et mécaniques.

Les volumes 4 et 5 abordent leurs versions synchrones basées sur la communication par cycle (SSRAM et SDRAM, la première lettre de leur acronyme signifiant *Synchronous*) et par paquets (RDRAM pour *Rambus DRAM* et SLDRAM pour *Synchronous-Link DRAM*).

Le volume 6 détaille la mémoire morte à semi-conducteurs et les mémoires non volatiles (NVM pour *Non Volatile Memory*) émergentes. La mémoire morte ou ROM (*Read-Only Memory*) est indispensable à l'ordinateur, qu'il soit à usage général ou spécialisé comme un système embarqué. Sa caractéristique principale est l'absence de volatilité, c'est-à-dire la perte de l'information avec l'arrêt de l'alimentation électrique. Elle stocke, en particulier, les programmes de test et d'amorce (*boot*) exécutés au démarrage et des informations de configuration qui, en leur absence, ferait que l'ordinateur ne pourrait démarrer. Les technologies actuelles de fabrication des circuits intégrés atteignent leurs limites. C'est ce que les spécialistes nomment le mur de briques rouges (*red brick wall*, voir la section 1.5 de (Darche 2021a)), qui est la limite physique de la finesse de la gravure des circuits intégrés. Les mémoires émergentes tentent de dépasser cette contrainte. Elles sont soit au stade industriel, soit au stade du prototype, au niveau de la puce ou de la cellule de mémorisation, voire du concept. Les modèles industriels présentés sont les mémoires ferroélectriques, magnétiques et à changement de phase.

Le volume 7 avec les mémoires vives spécialisées s'intéresse aux politiques de gestion (*management strategy or policy*) de la mémoire autre que l'accès aléatoire. Les politiques simples comme celle de « la donnée la première entrée est la première sortie » de la pile ou celle de « la donnée la dernière entrée est la première sortie » de la file y sont d'abord étudiées. Des modèles de mémoires dont l'interface ou la structure interne diffère de la DRAM classique mais qui restent à accès aléatoire, au moins pour un port de communication, sont ensuite présentées. Il s'agit des mémoires à pseudo-interface, à antémémoire et à applications spécifiques destinées en particulier à l'affichage.

Le volume 8 aborde la notion de cache. Ce mécanisme fonctionnellement transparent joue un rôle essentiel dans la hiérarchie des mémoires, car il permet d'améliorer de manière significative les performances des microprocesseurs modernes (c'est-à-dire à partir des années 1990) et des périphériques, principalement de mémorisation.

Le volume 9 détaille la notion de mémoire virtuelle. Elle a été proposée pour répondre à trois besoins essentiels de l'ordinateur. Il s'agit de la protection, de la translation de zone(s) mémoire(s) et de l'accroissement de la taille de l'espace de programmation. La réponse globale a été la correspondance d'adresse (*address mapping*) en séparant la mémoire logique de la mémoire physique grâce à des adressages distincts.

Les références dans le texte suivent la convention suivante : « voir section V5-2.3 », par exemple, signifie se référer à la section 2.3 du volume 5. Sans indication d'un numéro de volume, il s'agit d'une référence dans le volume courant.

La forme italique est utilisée pour des termes en langue étrangère comme l'anglais ou ancienne comme le latin.

Une organisation à plusieurs niveaux

Chaque volume de la série glisse progressivement du conceptuel vers l'implémentation matérielle. La pédagogie a été mon souci principal, sans pour autant négliger l'aspect formel. La lecture se veut à plusieurs niveaux. Chaque lecteur puisera les connaissances adaptées avant d'aborder la compréhension de certaines parties plus difficiles. Les connaissances, à quelques exceptions près, ont été présentées linéairement et de la manière la plus exhaustive possible. Des exemples concrets puisés dans les technologies anciennes et actuelles illustrent les concepts théoriques.

Si nécessaire, des exercices complètent l'apprentissage en étudiant de manière plus approfondie certains mécanismes. Chaque volume se termine par les références bibliographiques renvoyant aux articles de recherche, aux ouvrages ou aux brevets à

l'origine des concepts et ceux, plus récents, reflétant l'état de l'art. Ces références permettent au lecteur de trouver plus de renseignements ou d'informations théoriques. Une liste des acronymes employés et un index couvrant l'ensemble de l'ouvrage s'y trouvent aussi.

Cette série d'ouvrages consacrée à la mémorisation dans les ordinateurs est le fruit de plus de quarante ans de pérégrinations dans les mondes électronique, micro-électronique et informatique. J'espère qu'elle vous apportera une connaissance suffisante, à la fois pratique et théorique pour, ensuite, vous spécialiser dans l'un de ces domaines. Je vous souhaite pour ce premier ouvrage une agréable balade dans ces différents mondes...

REMARQUE. Il ne s'agit pas d'un ouvrage sur la technologie de fabrication des puces mémoires, bien que le sujet soit abordé. Il existe pour ce sujet des ouvrages spécialisés de référence comme (Prince 1996, 1999 ; Itoh 2001 ; Haraszti 2002 ; Sharma 2002a, 2002b ; Chen 2003 ; Redaelli et Pellizzer 2022 ; Yu 2022) ou, pour des familles particulières, (Brown et Brewer 1997 ; Keeth et Baker 2000 ; Campardo *et al.* 2005 ; Brewer et Gill 2008).

Comme il s'agit d'un ouvrage introductif au domaine de la mémorisation, des références de composants industriels de toute époque sont citées. Les noms des sociétés sont ceux d'origine, même si certaines ont fusionné. Cela permettra en particulier aux lecteurs de retrouver sur Internet les feuilles de caractéristiques des circuits intégrés cités et les documentations d'origine et de les étudier en relation avec cet ouvrage.

Les concepts présentés reposent sur les notions étudiées dans les ouvrages (Darche 2000, 2002, 2003, 2004, 2021a, 2021b, 2021c, 2021d, 2021e) qu'il est recommandé de lire au préalable.

Introduction

Ce volume est consacré à l'un des trois sous-ensembles du modèle de machine dite de « von Neumann » (von Neumann 1945) qu'est la mémoire¹. La fonction de mémorisation, essentielle au fonctionnement du processeur, s'est complexifiée progressivement pour répondre à une demande accrue de capacité, de sécurité ou de rapidité et, maintenant, de sobriété énergétique. Elle est omniprésente. On la retrouve dans la mémoire primaire et la cache bien évidemment, mais aussi dans le processeur et les interfaces d'E/S. La mémoire primaire est constituée de composants intégrés spécialisés dans la mémorisation, la plupart du temps soudés sur une ou deux faces d'un petit circuit imprimé (appelé « barrette » ou « rangée de mémoires » (*memory rank*)). La mémoire peut aussi être un bloc d'une puce plus complexe comme un microprocesseur (μ P ou MPU pour *MicroProcessor Unit*) ou un microcontrôleur (μ C ou MCU pour *MicroController Unit*). Il peut s'agir alors d'une banque de registres ou de zones mémoires intégrées de taille plus importante comme une cache² ou une mémoire primaire (*primary memory*) de petite taille. On l'appelle alors « mémoire embarquée » (*embedded memory*). On la retrouve aussi dans les contrôleurs d'E/S sous la forme de registres, d'une cache ou sous la forme d'un tampon qui gère le flux d'informations selon la politique « premier entré, premier sorti » ou FIFO (*First In, First Out*). Un périphérique peut aussi posséder une zone mémoire tampon FIFO ou de type cache et une zone de stockage. Citons l'imprimante ou l'unité de mémoire de masse à disque(s) dur(s) ou HDD (*Hard Disk (HD) Drive*) ou à mémoires à semi-conducteurs ou SSD (*Solid-State Disk*). Avec ce dernier exemple, le périphérique est

1. Les deux autres sous-ensembles sont les entrées-sorties (E/S) et le processeur, respectivement traités dans (Darche 2003) et (Darche 2021a, 2021b, 2021c, 2021d, 2021e).

2. L'origine de ce terme est française. Il qualifie un endroit secret de stockage, un synonyme de ce mot étant « cachette » qui sera réservé à la version parallèle. C'est pour cette raison que nous utiliserons le genre féminin pour ce nom dans cet ouvrage.

alors une mémoire dite de masse par la quantité d'informations stockées. Il appartient alors la catégorie des mémoires secondaires ou tertiaires.

La mémoire et son environnement, c'est-à-dire son contrôleur, s'il existe, et son bus de communication appelé aussi « canal », ont un impact direct sur les performances d'un ordinateur (puissance de calcul, consommation électrique, etc.) et son coût. Il est donc essentiel que l'électronicien et l'informaticien connaissent bien son fonctionnement et les technologies associées. L'électronicien s'en servira pour concevoir ce sous-ensemble et le contrôleur associé. L'informaticien, lui, optimisera au mieux ses programmes, en particulier dans le domaine du calcul scientifique.

Ce premier volume, prolégomènes du domaine, est organisé en six chapitres. Le premier débute par une présentation des caractéristiques principales d'une mémoire générique. La caractéristique principale est sans doute sa taille et l'unité de mesure standardisée associée est détaillée. La quête des chercheurs en architecture des ordinateurs et des industriels a été d'inventer des technologies de mémorisation fiables et suffisamment rapides pour un coût minimal ou économiquement acceptable. Les technologies anciennes et actuelles sont ensuite décrites. Pour terminer, le concept de hiérarchie des mémoires est étudié. Le second chapitre s'intéresse plus particulièrement aux modèles à semi-conducteurs (*solid-state memory*) à accès aléatoire. Il détaille leur organisation fonctionnelle interne avec les principaux blocs logiques. Le chapitre 3 complète le précédent avec des blocs spécifiques à certains modèles comme la boucle à verrouillage de phase ou PLL (*Phase-Locked Loop*). Le chapitre 4 étudie les organisations modulaires comme le bloc (ou sous-zone mémoire) et la banque et aborde les architectures multithreadées. L'avant-dernier chapitre se tourne vers les aspects externes pour augmenter le format de donnée et la capacité. Il s'intéresse aussi à l'entrelacement et au décodage des adresses ainsi qu'à l'interfaçage. Ce volume se termine par quelques notions complémentaires comme le cadrage, l'ordre de rangement des informations, la topographie mémoire et la détection et la correction d'erreur(s). L'annexe 1 est une introduction aux codes détecteurs/correcteurs d'erreurs.