

ORGANISATION D'UN ORDINATEUR

Les microprocesseurs sont utilisés avec d'autres composants LSI pour réaliser des ordinateurs. Le microprocesseur remplit toutes les fonctions de l'unité de traitement, alors que les autres composants fournissent la mémoire, les entrées/sorties, ainsi que les autres fonctions requises par le système.

Tous les ordinateurs d'usage général présentent la même organisation logique avec cinq unités fondamentales, comme le montre la fig. 1.

L'*unité centrale de traitement* regroupe deux modules :

- l'unité de commande *Control Unit* (CU),
- l'unité arithmétique et logique *Arithmetic and Logic Unit* (ALU).

L'ALU effectue les opérations arithmétiques et logiques sur les données qui la traversent, telles que les additions et les soustractions, ET logique, le OU logique et les opérations de décalage.

Le rôle principal de l'unité de commande est de chercher, de décoder et d'exécuter les instructions successivement d'un programme stocké dans la mémoire. Elle régit le séquençement de l'ensemble du système. Elle engendre et gère les signaux de commande nécessaires au cadencement des opérations, et commande la circulation des instructions du programme et des données à l'intérieur et à l'extérieur de l'unité arithmétique et logique (UAL). L'unité de commande est généralement reliée à l'UAL dont elle régit le fonctionnement.

Le module de mémoire sert à stocker les informations. D'un point de vue fonctionnel, la mémoire peut contenir deux types d'informations : les programmes et les données. Un programme est une suite d'instructions codées sous forme binaire de façon à ce quelles puissent résider dans une mémoire :

- le programme précise la suite d'instructions exécutées par l'ordinateur. Sous la supervision de l'unité de commande, chacune des instructions successives du programme est recherchée en mémoire puis déposée dans un registre spécial de l'unité de commande, où elle sera exécutée. A titre d'exemple, une instruction

type serait d'additionner les contenus de deux registres et de déposer le résultat de cette addition dans un troisième registre.

Les données stockées en mémoire sont traitées par l'UAL.

On distingue deux types de mémoires logiques, suivant les possibilités d'accès : les mémoires RAM (*Random Access Memory* = Mémoire à accès aléatoire) qui peuvent être lues ou écrites, et les mémoires ROM (*Read Only Memory* = Mémoire à lecture seule) qui ne peuvent être lues qu'une fois que les données y ont été inscrites. La mémoire ROM est une mémoire non volatile. Une RAM LSI standart est une mémoire volatile : son contenu est perdu lorsque l'alimentation est coupée.

Les deux autres modules de de la fig 1 sont *le module d'entrée et le module de sortie*. Leur rôle est d'assurer la communication avec le monde extérieur. *Le module d'entrée* fournit l'information à l'UAL ou à la mémoire. Les dispositifs d'entrées courants sont le clavier ou le capteur, tels que le capteur de température, le détecteur de présence, ou le capteur de pression. *Le module de sortie* affiche les données en provenance de l'UAL. Les circuits de sortie courants sont l'imprimante, le CRT..., ou l'afficheur à cristaux liquides.

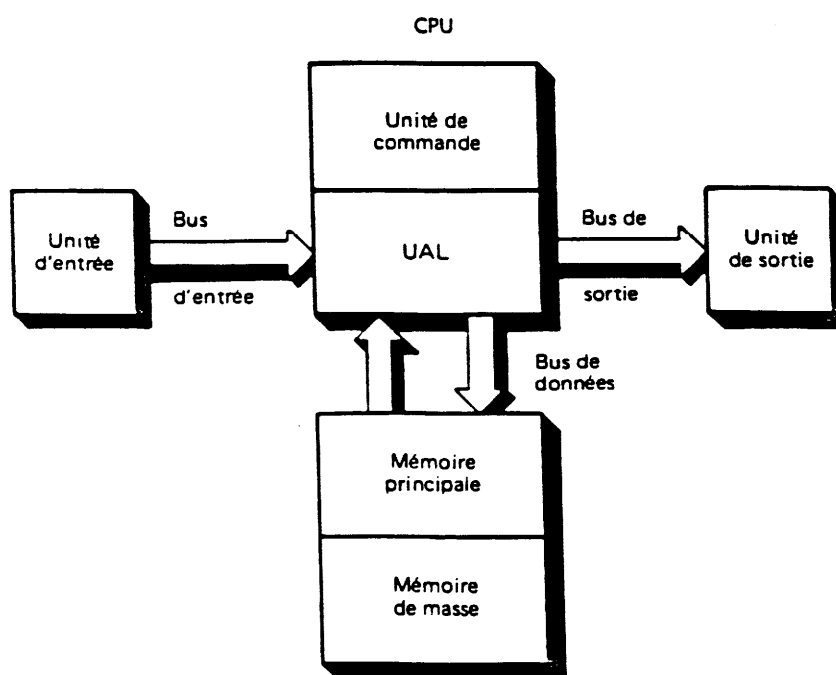


Figure 1

MACHINE VON NEUMANN

**1945: CONCEPTION DE "ELECTRONIC DISCRET VARIABLE AUTOMATIC
COMPUTER " : EDVAC**

FABRICATION EN 1950

1947: PUBLICATION DU RAPPORT SUR 'L'ARCHITECTURE VON NEUMANN .

**INTRODUCTION DE LA NOTION DE PROGRAMME -> STOCKER DANS UNE
MEMOIRE**

ARCHITECTURE VON NEUMANN

AMELIORATION DE LA MACHINE DE TURING

FORMAT DE TRAVAIL

ACCES ALEATOIRE

DECOMPOSITION D'UNE INSTRUCTION EN PLUSIEURS PHASES

TEST CONDITIONNEL SUR UN REGISTRE PARTICULIER

LANGAGE DE PLUS HAUT NIVEAU -> COMPILATEUR

ARCHITECTURE VON NEUMANN

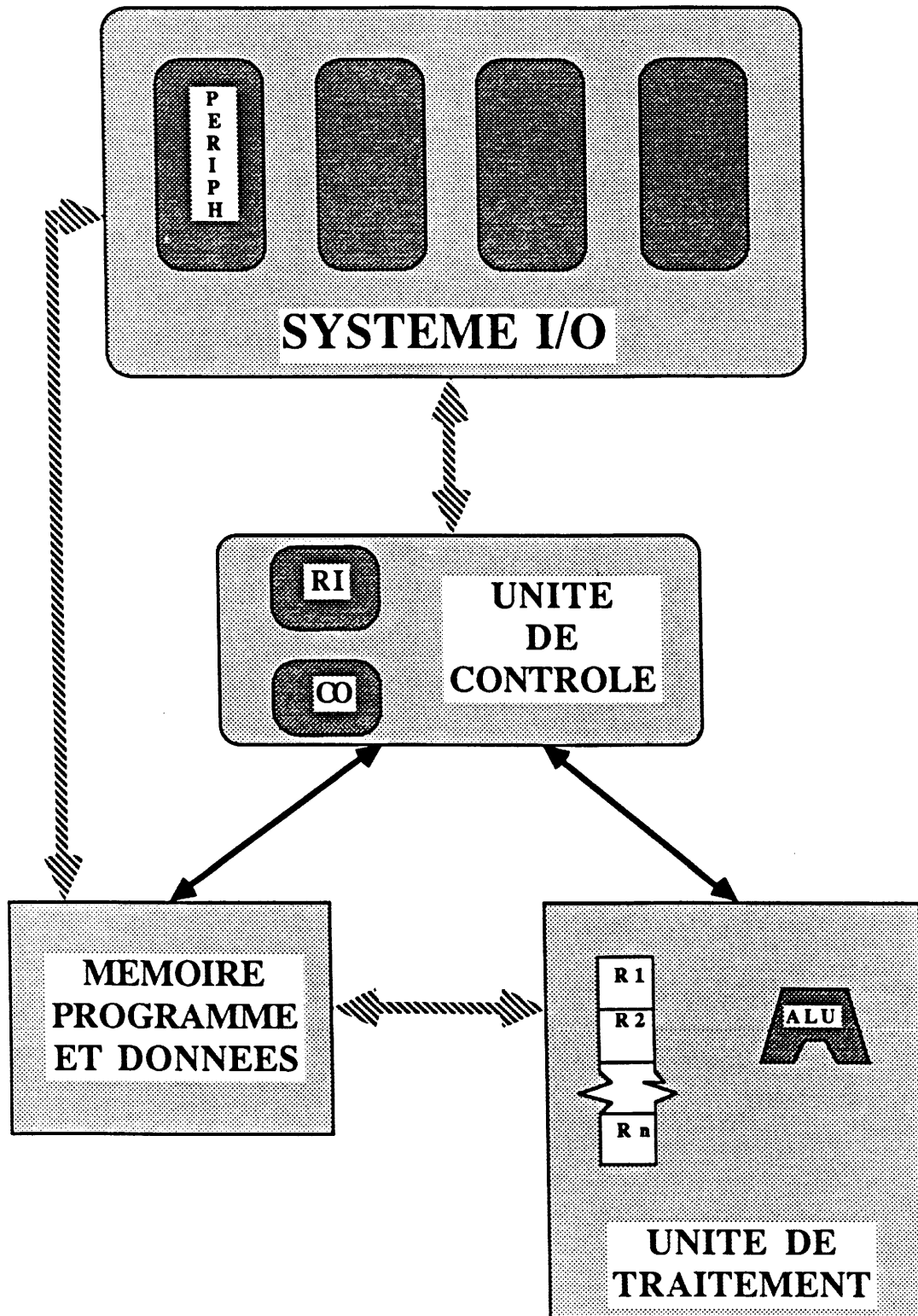
QUATRE MODULES SONT INTERCONNECTES

UNITE MEMOIRE

UNITE DE TRAITEMENT

UNITE DE CONTROLE

SYSTEME ENTREE/SORTIE



MACHINE VON NEUMANN

CRITIQUES :

MELANGE DONNEES / INSTRUCTIONS

ACCES SEQUENTIEL

OCCUPATION DU PROCESSEUR

ARCHITECTURE HARVARD :

APPARUE DEPUIS 1980

SEPARATION PHYSIQUE MEMOIRE PROGRAMME / MEMOIRE DONNEES

AVANTAGES : ELIMINE LE GOULET D'ETRANGLEMENT

PB : STRUCTURE COMPLEXE ; AUGMENTATION DE PLACE

PROBLEMES INADAPTES A L'ARCHITECTURE VON NEUMANN

TEMPS REEL

IMAGERIE :

TRAITEMENT DE L'IMAGE

SYNTHESE D'IMAGE

ANALYSE DE L'IMAGE ; RECONNAISSANCE DE FORMES

MODELISATION :

PHENOMENES ATMOSPHERIQUES

METEOROLOGIE

MODELISATION DE PIECES ET STRUCTURES :

POURQUOI DES GIGAFLOPS ?

• Prévision météorologique :

prévision à 24 H sur l'ensemble du Globe ("grossière")
→ $2000 \cdot 10^9$ OPÉRATIONS par point moyen de calcul

• Construction Aéronautique

"Soufflerie" classique → Soufflerie Numérique
étude en 3D des profils d'avions, turbulences ...

$5000 \cdot 10^9$ opérations

Navette spatiale: $10\,000 \cdot 10^9$ opérations par configurations

→ Machine stratégique pour l'Aéronautique

$[10^{13}$ opérations flottantes 64 bits avec 1 MFLOPS en moyenne
(GOULD NPA 1,7)
⇒ $\approx 3000 \text{ H}$ → 3 mois: 24/24 H

100 MFLOPS en moyenne

⇒ 30H ...

Historiques des processeurs Intel

Intel Corporation, société fondée en 1968, a conçu le premiers microprocesseurs universels en 1971.

Le 4004

Le 4004 était prévu pour de l'arithmétique BCD (Binaire Codé Décimal), il ne comportait que 2250 transistors en technologie PMOS mais remplaçait six circuits spécialisés qui n'auraient été utilisables que par un seul client (car non programmables).

Le 8008

En 1972 est sorti le premier microprocesseur 8 bits : le 8008, avec un jeu d'instructions différent du 4004. Il était initialement prévu pour piloter un terminal CRT, Intel avait rajouté par la suite des instructions pour en faire un processeur multiusages.

Le 8008, avec seulement 3300 transistors PMOS, est l'ancêtre de la famille des 80x86, on y trouve la possibilité de faire :

- des calculs arithmétiques 8 bits (non-signées),
- 4 indicateurs d'état (signe, zéro, parité, retenue),
- un accumulateur A et 6 autres registres de 8 bits,
- une pile matérielle pour 8 adresses de retour,
- une capacité d'adressage de 16 kilo-octets (PC = 14 bits).
- 66 codes d'instructions,
- adressages indexé immédiat.

Le 8080

En 1974, le 8080 apporte surtout une évolution technologique avec *le passage du PMOS au NMOS* : avec 4500 transistors NMOS, il multiplie par 10 la puissance de traitement du 8008 en gardant une compatibilité source, et devient rapidement un standard de l'industrie.

Le 8080 offre un *espace d'adressage de 64 ko* (bus d'adresse de 16 bits). De nouveaux modes d'adressage sont possibles tel que *l'adressage direct* pour charger l'accumulateur, et une pile de taille illimitée est pointée par *SP (Stack Pointer)*. Il devient possible de *calculer en BCD*, avec l'introduction d'un indicateur d'état détectant la retenue entre quartets, et le *mécanisme d'interruptions* original du 8008 devient praticable grâce à la possibilité de sauver les indicateurs d'état sur la pile et d'autoriser ou inhiber les interruptions.

Le 8085

Le 8085 est un 8080 avec un contrôleur d'interruptions (8228) et d'entrée-sortie série intégré, et un circuit d'horloge (8224). Seulement deux instructions le différencient du 8080 (pour à la fois accéder au masque d'interruption et gérer l'entrée-sortie série).

La généalogie Intel s'est poursuivie à partir du 8080 et non du 8085, qui peut être considéré comme un microcontrôleur. D'autres microcontrôleurs se sont basés sur le 8080, intégrant parfois de la mémoire morte ou des entrées-sorties; par exemple, on trouve encore aujourd'hui dans le clavier du PC un 8048.

Le 8080 a eu tellement de succès que des processeurs compatibles concurrents sont apparus, comme le Z80 de Zilog en 76, avec un jeu d'instructions plus homogène et très nettement étendu (manipulations de bits, de chaînes...).

Le 8086

Sorti en 1978 (mais la conception avait commencé en 75), le 8086 est un processeur de 29000 transistors en technologie CMOS (moindre consommation). Cette fois-ci, Intel n'était pas le premier à sortir un microprocesseur 16 bits : en 76 étaient apparus des processeurs qui n'ont pas laissé de grandes traces dans l'Histoire (TMS9900 de Texas Instrument, 9440 de Fairchild, LSI11 de DEC...).

La compatibilité binaire avec le 8080 disparaît mais un moyen mécanique permet de récupérer les sources (deux instructions ont d'ailleurs été introduites afin de traduire plus efficacement la sauvegarde et la restauration du PSW du 8080 sur la pile : LAHF et SAHF).

L'objectif du 8086 était d'apporter des extensions symétriques au 8080 (arithmétique signée 8 et 16 bits, chaînes, programmes relogeables et réentrants) tout en multipliant les performances par un facteur 10.

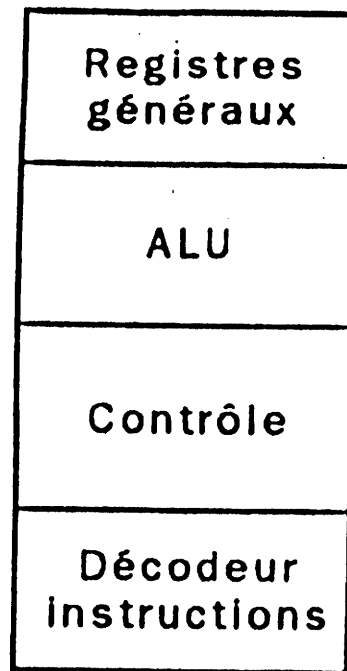
La notion d'accumulateur disparaît un peu mais les registres restent souvent spécialisés : BX comme pointeur de base en mémoire, CX comme compteur, DX comme données et AX comme accumulateur pour certaines instructions (multiplication, division, entrées-sorties), SI et DI comme index source et destination, SP comme pointeur de pile et BP comme pointeur de base sur la pile. Les modes d'adressages deviennent puissants (possibilité d'ajouter deux registres base et index, et un déplacement) mais lents.

La notion de pipelining empruntée aux gros calculateurs fait aussi son apparition avec le prefetch (deux unités travaillent en parallèle : l'Execution Unit et la Bus Interface Unit qui s'occupe de tous les accès mémoires et anticipe le chargement des instructions).

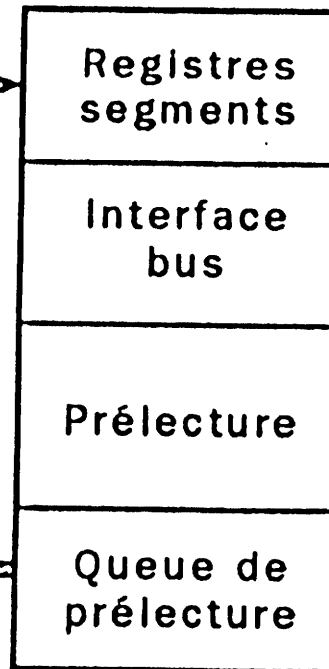
En 79 sort une version bus de donnée 8 bits : le 8088 qui sera incorporé dans le PC d'IBM.

La structure interne du 8086/8088

Unité d'exécution



Unité d'Interfaçage bus



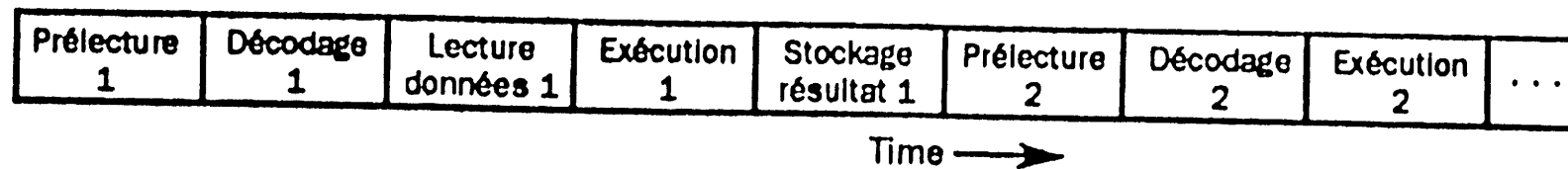
16
Requêtes de
données et
opérandes

Prélecture
d'Instructions

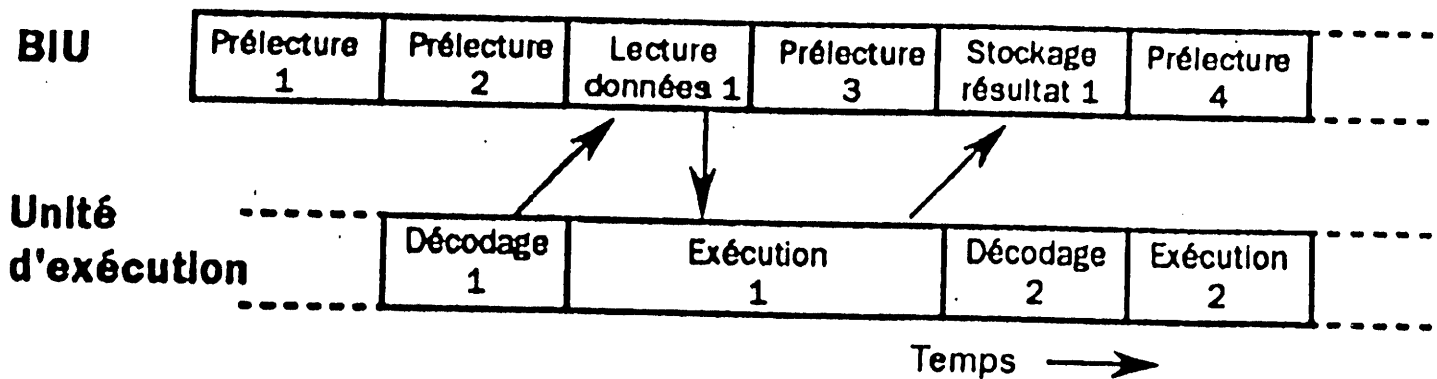
Bus
adresses/données
multiplexé
8086 = 16 bits
8088 = 8 bits

Traitements séquentiels et avec pipelining

Traitements en série



Traitements en parallèle



Le 80186

Le 80186 est un microcontrôleur : il intègre le 8086 avec une unité de DMA, un contrôleur d'interruptions, une horloge et des compteurs/temporisateurs, et un décodage d'adresse pour sélectionner des boîtiers mémoires. Le calcul des adressages mémoire est câblé, donc très rapide, et dix nouvelles instructions voient le jour.

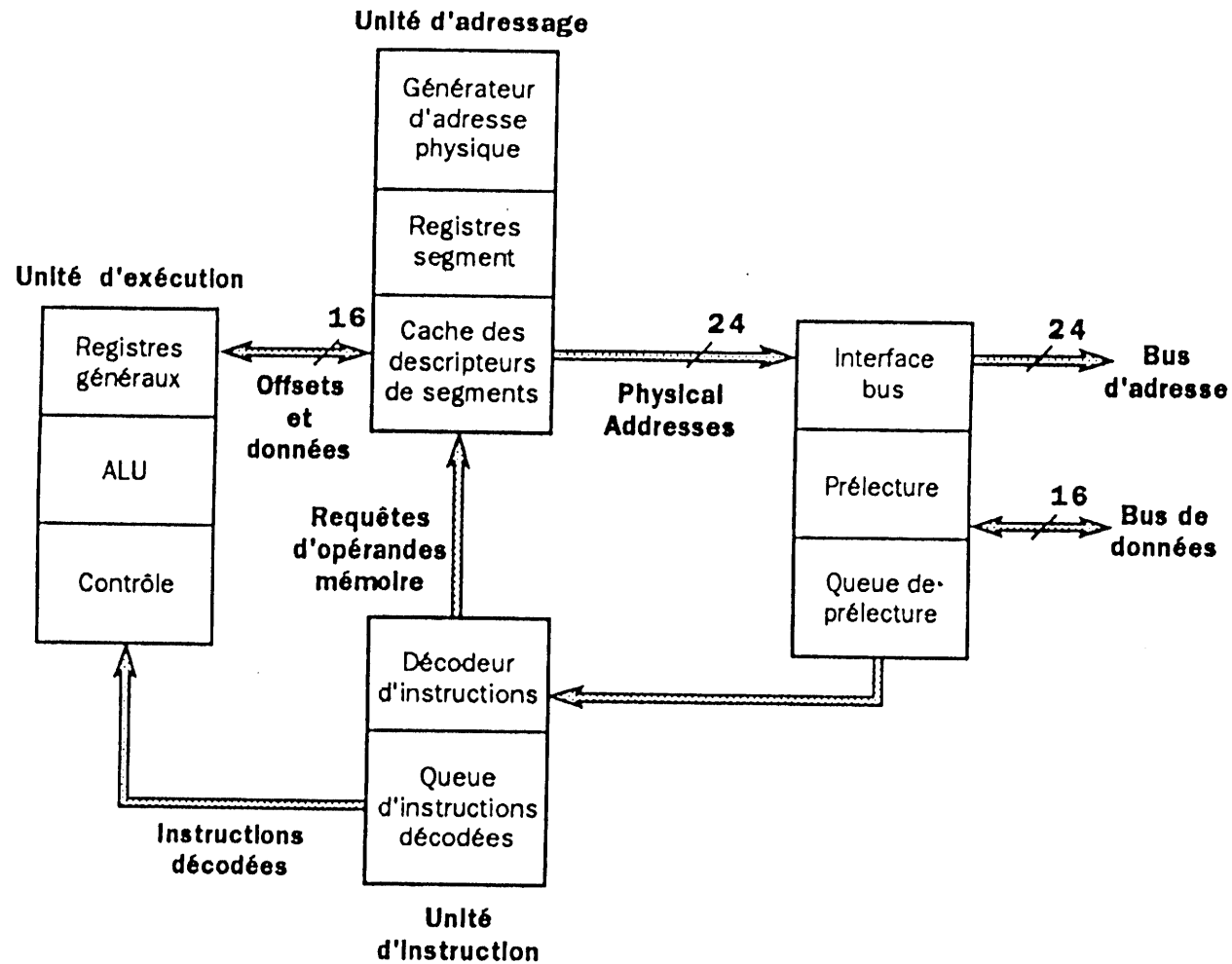
Le 80286

Le 80286 élimine le multiplexage du 8086 avec un pavé de 68 broches et introduit un adressage virtuel segmenté tandis que l'espace d'adressage physique passe à 16 Mo.

Deux modes donnent des comportements différents : le mode réel pour garder une compatibilité binaire avec le 8086, et le mode protégé où les segments deviennent des sélecteurs dans des tables de descripteurs (jusqu'à 8192 par table), avec une capacité d'adressage de 1 Go pour chaque tâche. Les tâches sont gérées directement par le processeur, ainsi que les protections des segments avec 4 niveaux de privilèges.

Le 286 accentue le pipelining en permettant à quatre unités de travailler en parallèle : la Bus Unit (interface avec le bus), l'Instruction Unit (décodage des instructions), l'Execution Unit et l'Address Unit (génération des adresses physiques à partir des adresses logiques segmentées).

La structure interne du 80286



Le 80386

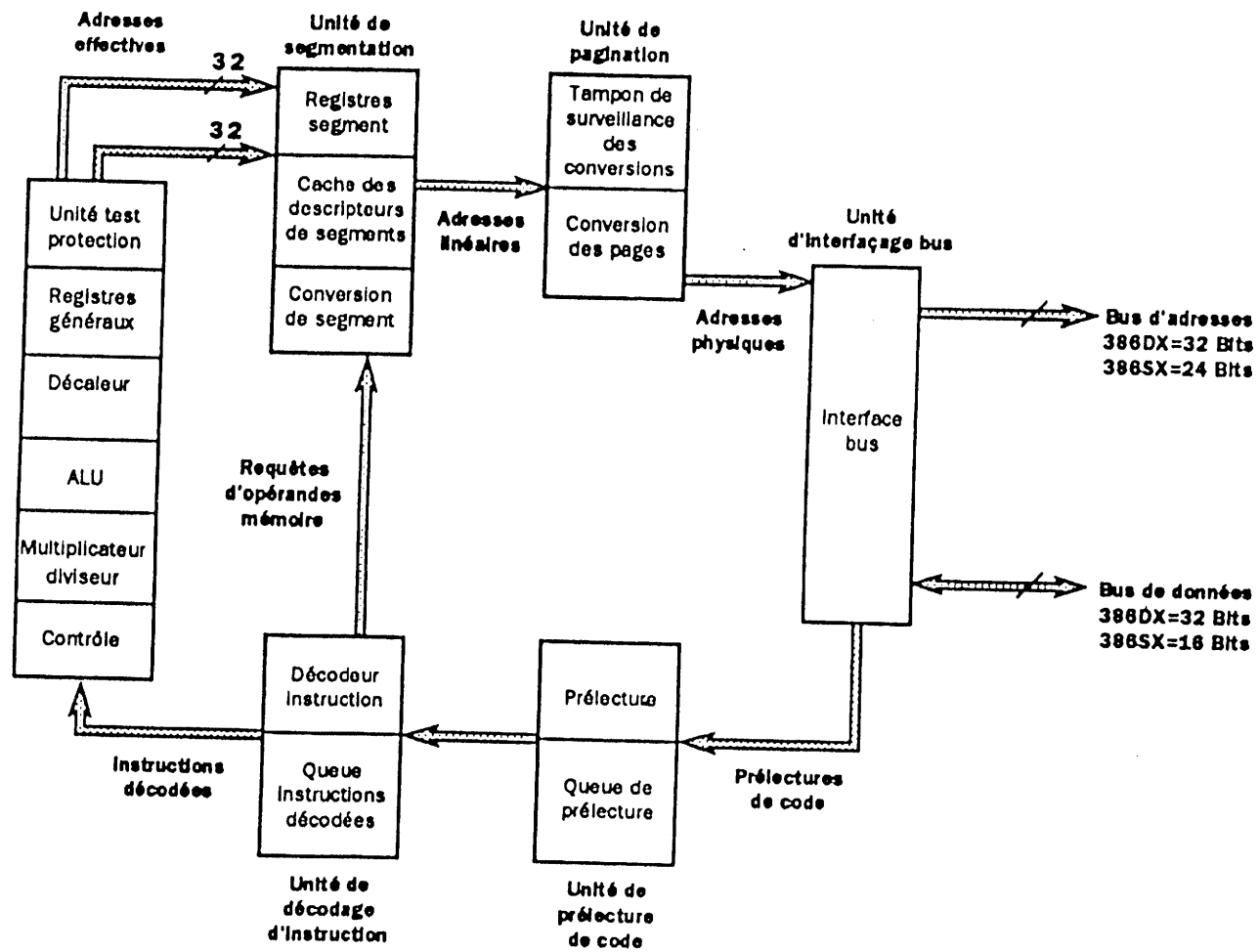
Le 80386 est le premier processeur 32 bits de la famille 80x86 (Intel avait lancé auparavant en 83 l'iAPX 432 qui offrait un espace virtuel de 1 Téra octet et un jeu d'instructions orienté langages de haut niveau, tel que ADA, mais il n'était pas compatible avec la famille x86 !). Il reste compatible avec le 8086 en proposant toujours un mode réel, et aussi avec le 286 avec le mode protégé qui permet d'exécuter aussi bien des tâches 16 bits (286) que des tâches 32 bits (386). Un autre type de tâche est par ailleurs introduit : les tâches virtuelles 86, qui permettent d'exécuter les programmes du 8086 en environnement protégé.

Le *pipelining* est encore renforcé, et la plupart des instructions sont maintenant exécutées en 2 à 5 cycles d'horloge (suivant qu'un accès mémoire est nécessaire ou non).

Le mécanisme de pagination est aussi intégré au processeur : le gestionnaire de mémoire est donc capable à la fois de segmenter et de paginer (l'espace d'adressage virtuel par tâche passe à 64 Téra-octets, pour un espace mémoire physique extensible à 4 Go).

Les adressages mémoires deviennent très symétriques et très puissants : un facteur d'échelle est introduit pour l'accès à des tableaux de mots de 16, 32 ou 64 bits. Multiplication, division et décalages gagnent en efficacité et de nombreuses instructions de manipulation de bits apparaissent.

La structure interne du 80386DX/SX

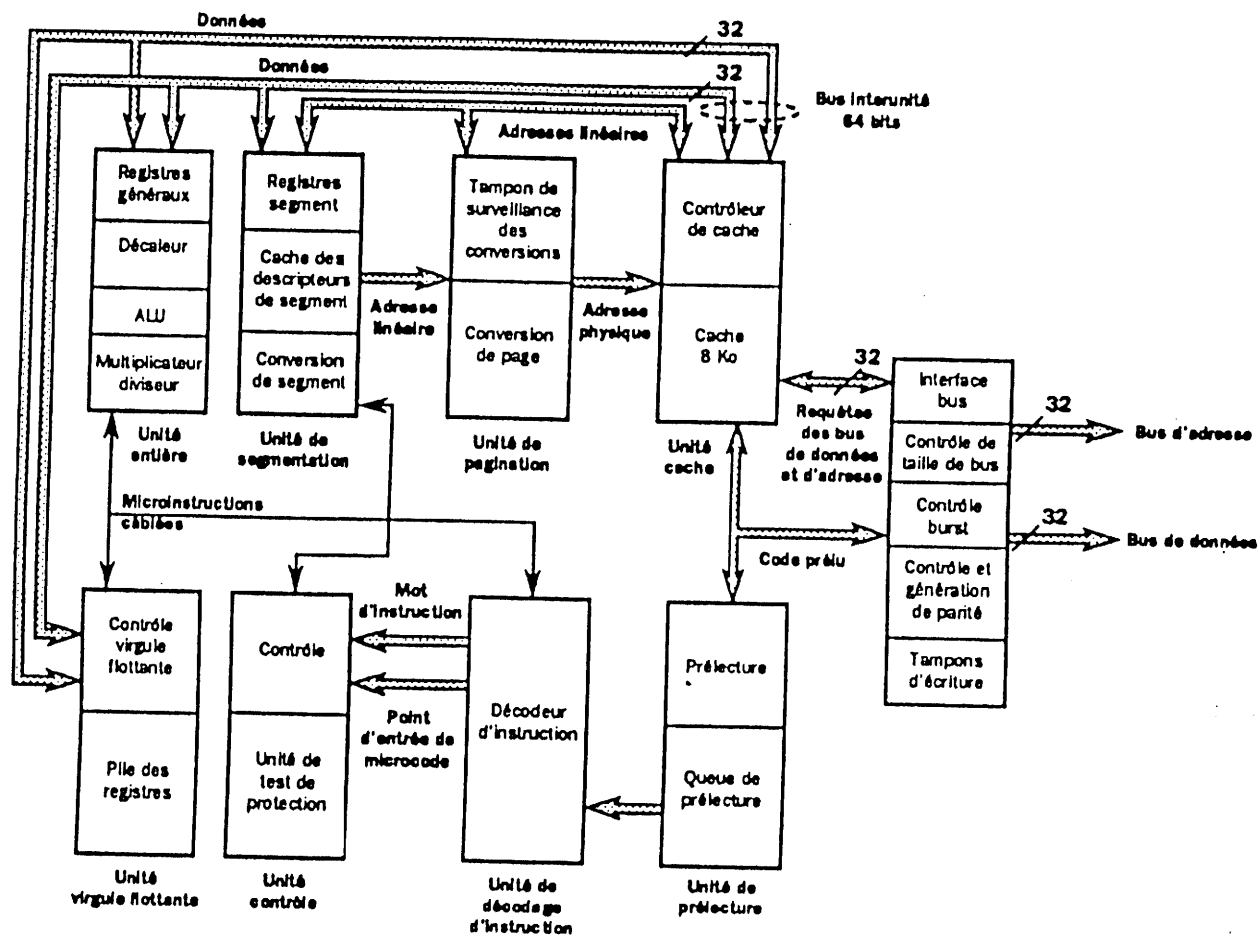


Quelques idées sont empruntées au concept RISC : un cache de 8 Ko est intégré (commun aux instructions et aux données) et la tendance au pipelining s'accentue encore, avec un décodage en deux étapes).

Le coprocesseur arithmétique habituel du 386 (le 80387) est aussi intégré dans la puce 486 : le coprocesseur a ainsi accès immédiatement aux opérandes et le calcul en virgule flottant est largement accéléré.

Avec une optimisation des instructions et le cache interne, les instructions les plus courantes s'exécutent en un seul cycle d'horloge (on ne compte ici que la durée de l'étage d'exécution en espérant que le pipeline reste rempli).

La structure interne du 80486



Le Pentium

La tendance au RISC est de plus en plus marquée : pratiquement tout le silicium de la puce est consacré à des caches (instructions et données disposent maintenant de caches séparés afin que le chargement des instructions puissent se faire en même temps que l'accès aux données) et à des unités d'exécution pipelinées (3% seulement est consacré à l'exécution d'instructions complexes microcodées).

Le processeur est d'autre part superscalaire, avec deux unités entières qui peuvent travailler en parallèle évidemment, il faut pouvoir récupérer les opérandes mémoires requis par les deux unités, et à cet effet, le bus de données du processeur passe à 64 bits.

Les deux unités entières sont pipelinées comme sur le 486 mais pour fonctionner en parallèle, il faut deux instructions simples consécutives dans le code.

Le Pentium Pro

Le Pentium Pro est la sixième génération des processeurs 80x86. Il accentue l'orientation prise par le Pentium en direction des dernières techniques des processeurs Risc: superscalaire, avec de larges caches internes, et surtout ce que Intel appelle exécution dynamique des instructions, la prédiction dynamique des branchements, l'exécution non séquentielle et l'exécution spéculative.

Les instructions de la famille 80x86 sont maintenant décodées en opérations simples Risc (qu'Intel appelle micro-opérations). Ces instructions Risc sont plus facilement distribuées aux différentes unités d'exécution (deux entières, une de chargement, une de déchargement et une flottante), c'est ce qu'Intel appelle exécution dynamique.

Le Pentium Pro possède un super-pipeline de 14 étages,

Le premier étage calcule la prochaine adresse du pointeur d'instructions (en tenant compte des interruptions ou des branchements).

Les trois étages suivants permettent au Pentium Pro de charger deux lignes de cache de 32 octets chacune, de marquer les frontières des instructions et d'envoyer 16 octets alignés au décodeur.

Les étages suivants convertissent les instructions en micro-opérations grâce à 3 décodeurs travaillant en parallèle: 2 pour les instructions 80x86 simples qui, lorsqu'ils trouvent une opération complexe, la passent au troisième (parfois, l'instruction est encore trop complexe pour lui, alors elle est envoyée à un séquenceur de micro-code, comme dans les premiers 8086 !) Au final, on se retrouve avec des micro-opérations triadiques.

Les étages 7 et 8 préparent les micro-instructions pour une exécution superscalaire.

L'étage 7 renomme les registres logiques (les 8 registres entiers classiques et les 8 flottants) en registres physiques (il y en a 40 de plus), ce qui permet l'exécution spéculative.

A l'étage 8 sont rajoutés des informations d'état (état d'avancement de l'instruction, unités d'exécution qui peuvent traiter l'instruction...), ce qui porte la taille des micro-opérations à 118 bits, et ces dernières sont alors enfilées dans un buffer circulaire (qu'Intel appelle buffer de réordonnancement) jusqu'à être complètement exécutées et retirées.

Les étages 9 et 10 permettent au Pentium Pro d'exécuter les micro-instructions du buffer dans le désordre.

Tant qu'il n'y a pas de dépendances, le cpu peut continuer à prédire les branchements et exécuter spéculativement les instructions suivantes (jusqu'à 20 ou 30 après le pointeur d'instructions !).

L'exécution non séquentielle est gérée par une unité (station de réservation), qui distribue les micro-opérations aux 5 unités d'exécution (la vitesse normale semble être de 3 micro-opérations par cycle) aux conditions qu'elles ne soient pas bloquées par une dépendance, que les opérandes soient prêts et que l'unité d'exécution qui pourrait traiter le type d'opération soit disponible.

L'étage suivant (11) est l'exécution proprement dite des micro-opérations, mais peut se poursuivre pour une micro-opération assez complexe (flottante par exemple, la plupart des entières ne demandent qu'un cycle, comme les chargements, et les déchargements qui en prendraient deux sinon ont été découpés en deux micro-opérations pour pouvoir être recouvert par d'autres opérations). Une fois exécutée, la micro-opération est replacée dans le buffer avec un bit signalant qu'elle est terminée et y attend d'en être retirée. L'unité de retirement vérifie que les micro-opérations sont retirées dans l'ordre, et fait attention aux interruptions, aux exceptions, aux branches mal prédites, et si tout va bien, mémorise les résultats.

C'est là où les registres physiques sont réassignés aux registres logiques (et si un registre logique modifié conditionne l'état d'autres micro-opérations dans le buffer, elles sont mises à jour). Ces étapes représentent les étages 12, 13 et 14 du pipeline (mais l'étage 12 est aussi celui de l'exécution des micro-opérations complexes, ce qui fait que le retirement prend deux cycles).