

LES PROCESSEURS

Introduction

Dans le chapitre 5, nous avons vu les processeurs selon leur type (famille, fabricant). Ce chapitre va expliquer l'architecture interne de ces processeurs et leurs performances et même (je suis optimiste) de suivre les évolutions futures des processeurs modernes.

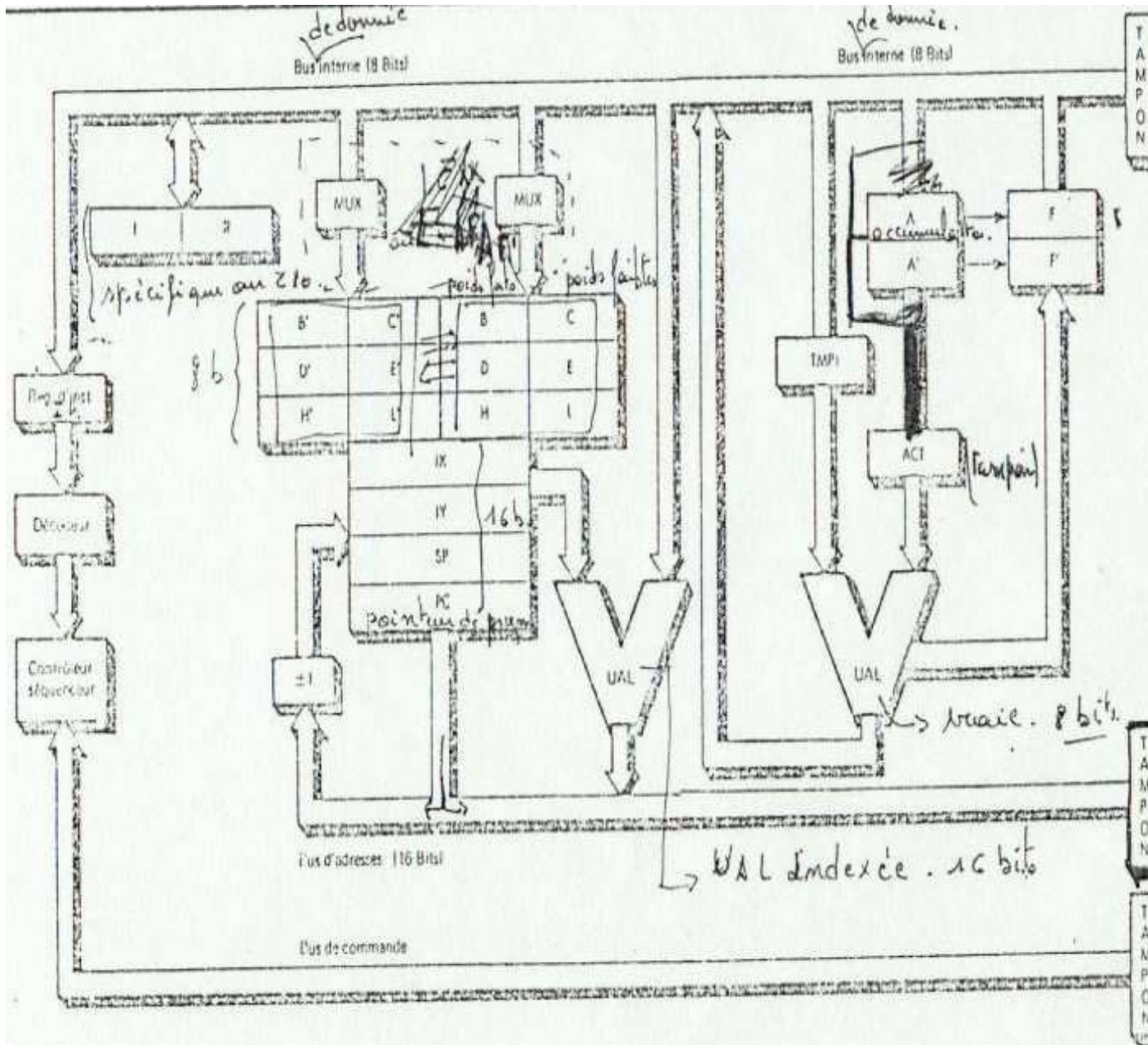
Les renseignements proviennent des notes techniques internes des fabricants de microprocesseurs, de quelques uns de mes cours en graduat chez Monsieur TANT (même s'ils remontent) et de synthèses de quelques revues. Ces schémas sont difficiles à trouver. Tous les "processeurs PC" utilisent les mêmes instructions que les premiers processeurs 8086 d'INTEL (mais souvent le système d'exploitation n'est plus compatible avec ces premiers programmes). Les instructions qu'un processeur est capable de lire sont programmées en assembleur. Tous les systèmes d'exploitation et programmes sont recodés à partir de langages évolués (C, pascal, Visual Basic, ...) vers cet assembleur. Si un processeur actuel utilise le même assembleur qu'un processeur de la première génération n'est pas limitatif. En cours d'évolution, certaines instructions ont été ajoutées pour permettre des décodages d'adresses sur une plage supérieure (386) ou pour des instructions multimédia (Pentium MMX, Pentium III, K6-2 et K6-3).

L'obligation de garder les instructions de base du 8086 pour les processeurs suivant pose de sérieux problèmes pour la conception de nouveaux microprocesseurs compatibles X86. En effet, les instructions du 8088-8086 ne sont pas faciles. Codées entre 1 et 5 bytes, elles ne permettent pas de prédire l'espace utilisé par l'instruction suivante. Suivant un codage à 8 bits des instructions, une instruction peut prendre plusieurs bytes. Par exemple, un chargement d'instruction dans la mémoire peut être de type STA 00 H0. Pour cette instruction de chargement d'une instruction de la mémoire (adresse 00 h0, codée sur 16 bits) vers l'accumulateur (une zone mémoire du processeur), l'instruction prend trois byte de données. Les processeurs 8086-8088 sont de types CISC (Complex Instruction Set Computer). Les premiers concepteurs de processeurs rajoutaient le plus d'instructions possibles pour permettre à l'utilisateur de peaufiner ses programmes. Néanmoins, ces multiples instructions ralentissent le fonctionnement des logiciels et sont peu utilisées en pratique. Actuellement, on utilise plutôt des processeurs de type RISC (Reduced Instruction Set Computer). Le nombre d'instructions est réduit, mais plus rapidement exécutées. Le fonctionnement actuel charge des instructions CISC 8086-8088 pour les recoder en instructions RISC plus rapides.

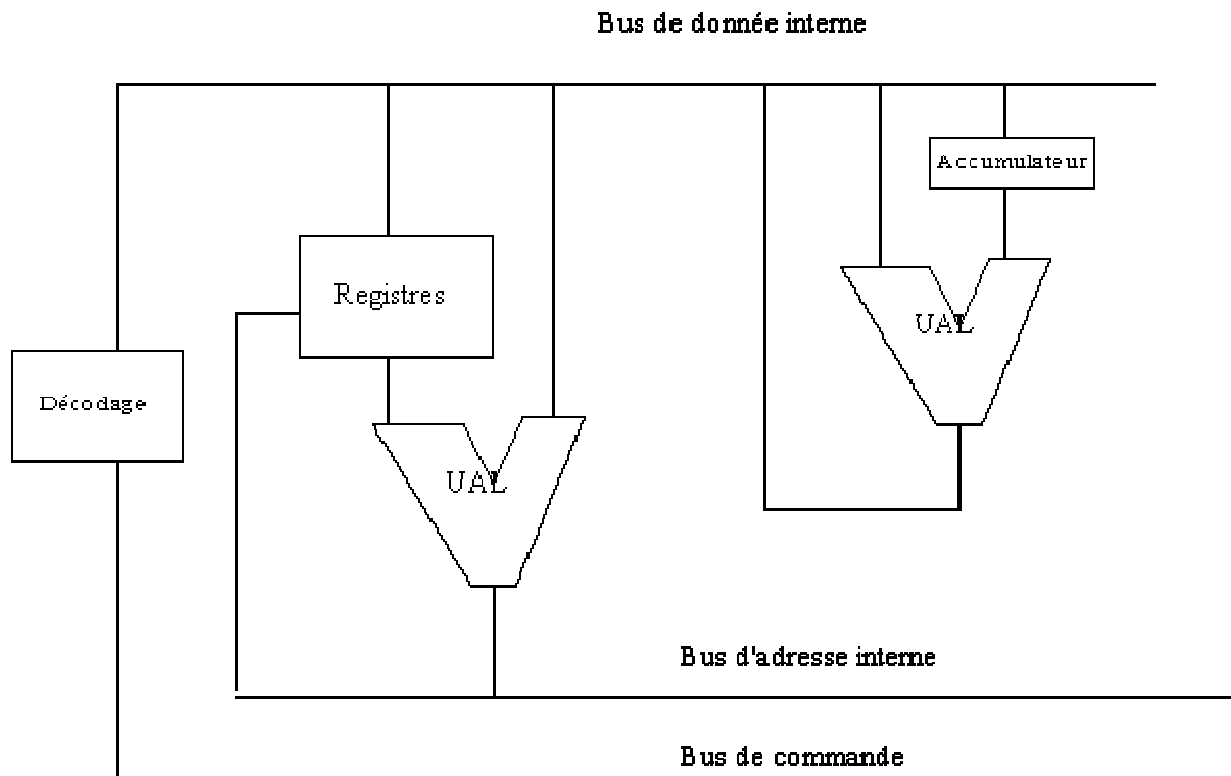
7.2 Structure interne d'un processeur de base.

Reprenons notre processeur Z80 (ou 6500, 6800 de structure comparable) de départ et analysons son fonctionnement. Sans rentrer dans les détails, distinguons les différentes parties: au dessus, un bus interne de donnée, en dessous un bus interne de commande et un bus interne d'adresse. Connaissant les structures des systèmes à microprocesseurs, ceci ne devrait pas nous étonner. Dans la partie de droite, une UAL Cette UAL est une unité de calcul, elle est connectée à une mémoire tampon (l'accumulateur) et au travers d'un bus à un ensemble de mémoire appelés registres. Dans le Z80, ces registres sont notés B, C, D, A', B', C', D' sont spécifiques au Z80, mais existent également dans les 8088. Ceci permet de

travailler avec deux zones mémoires que l'on fait basculer. L'UAL de gauche permet de faire des calculs d'adresses.



Le schéma des processeurs de cette génération reprenait un schémas suivant:



Cette structure charge un code d'instruction à la fois, soit pour une instruction de type LDA XX XX (chargement dans l'accumulateur de la donnée contenu à l'adresse mémoire XX XX), trois cycles d'instructions (une pour LDA et 2 pour l'adresse). Cette condition n'est remplie que si le processeur charge une instruction par cycle d'horloge, mais c'est le cas dans un système INTEL. Comme les instructions des nouveaux Processeur INTEL compatibles sont compatibles avec les 8088, les tailles des registres le sont aussi. Il est donc impossible d'augmenter la taille de ces mémoires pour augmenter la vitesse d'exécution.

Quelles possibilités d'augmenter les performances sans modifier la structure?

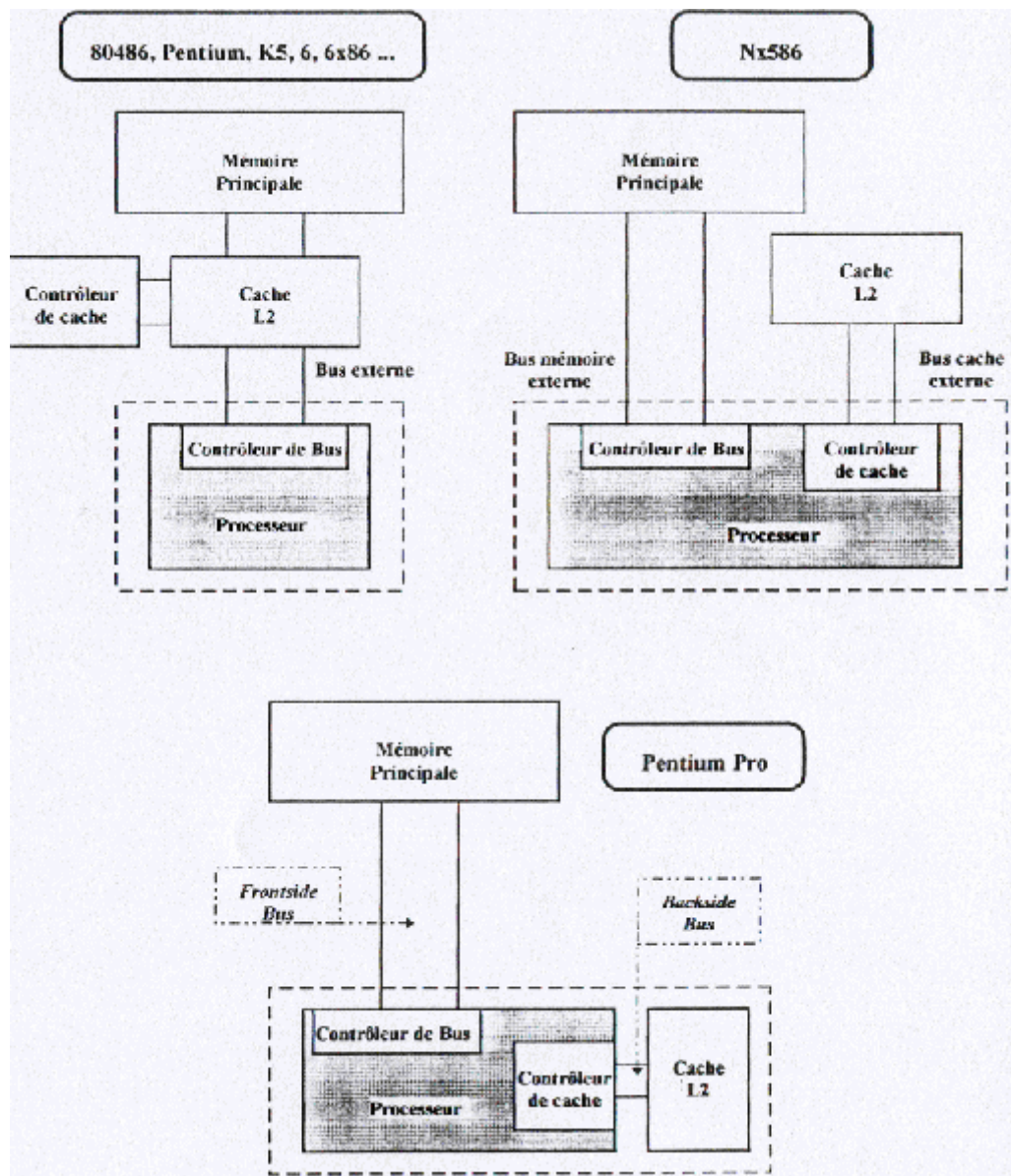
- . exécuter plusieurs instructions en même temps, mais cela est difficile puisque certaines instructions sont conditionnelles.
- . faire passer via un décodeur les instructions de type CISC en RISC plus rapide.
- . S'assurer que le processeur n'attend jamais d'instructions, qu'elles soient directement accessibles.
- . augmenter l'intégration en diminuant la taille des transistors. Ceci augmente la vitesse de transfert entre les différentes parties et réduit l'échauffement.

Tous ces trucs sont actuellement utilisés, tant par INTEL, que par AMD et Cyrix (VIA).

7.3. Utilisation d'une mémoire cache

En insérant une mémoire tampon rapide entre le processeur et la mémoire (plus lente que le processeur), on permet au processeur d'accéder rapidement à des données et

instructions qu'il utilise souvent. On distingue un cache interne L1 (implanté dans le processeur) et un cache L2 (extérieur au processeur, même s'il est implanté dans le même boîtier, cas des Pentium II, K6-2 et K6-3).



Les différentes architectures de cache

Selon le modèle de processeur, le cache L2 est géré différemment. Dans le cas des Pentium Pro et Pentium II, le cache L1 n'est plus en tampon direct entre la mémoire et le processeur, mais directement géré par le processeur. Ceci explique que le cache est implanté dans le même boîtier pour ces processeurs.

A titre d'exemple, un CELERON 300 sans cache L2 est 40 % inférieur en performance avec un Pentium II de même fréquence et de structure identique.

De plus, les processeurs actuels utilisent 2 caches L1 internes: un pour les données, un pour les instructions. Plus la taille du cache augmente, plus les difficultés à gérer ce cache augmentent, mais plus la vitesse augmente.

7.4. Passage du CISC au RISC.

Une autre solution pour augmenter les performances est de faire passer la structure des instructions du CISC (instruction 8088) en RISC. Chaque firme utilise en interne ses propres instructions, meilleur est ce transfert, meilleures sont les performances. Ceci explique que la vitesse effective d'un processeur n'est pas identique pour 2 processeurs de même fréquence, mais de structure différente.

Avantage des processeurs CISC:

. Instructions plus proches d'un langage de haut niveau

- Programmation plus compacte
- Ecriture plus rapide et plus élégante des applications
- Moins d'occupation mémoire des programmes
- Exécution nécessite moins d'octets mémoire

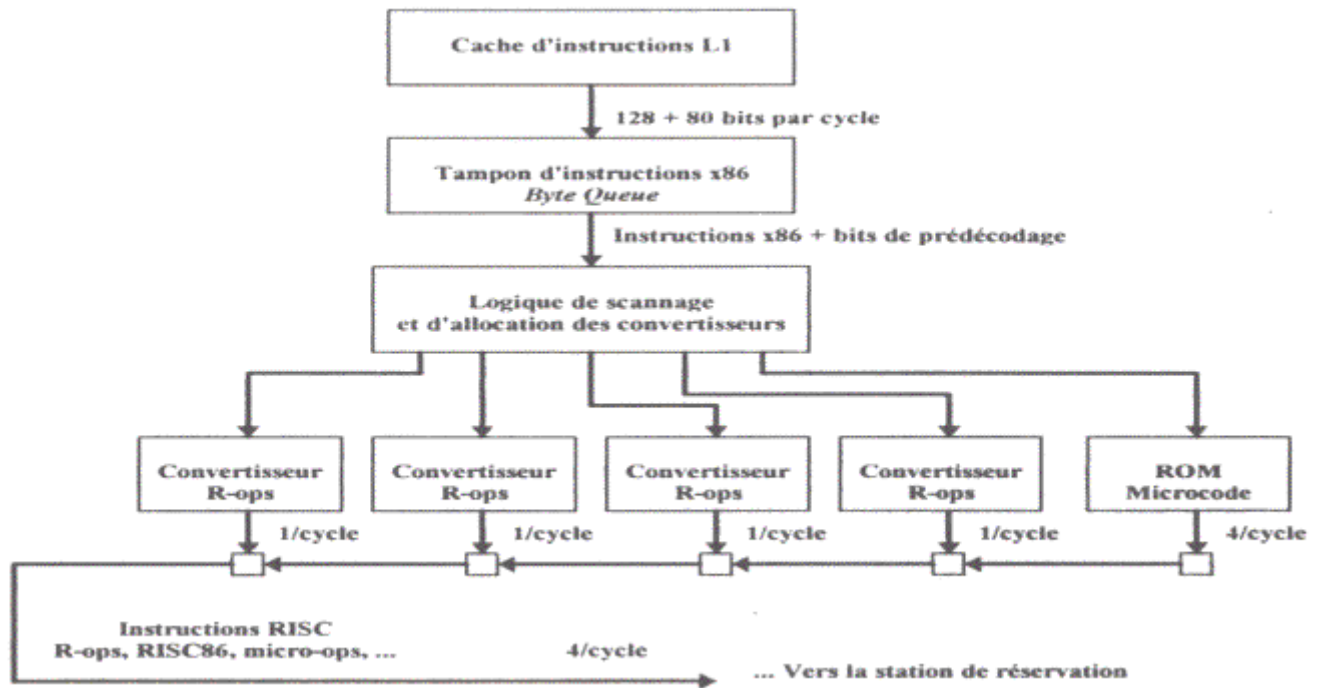
Inconvénient des processeurs CISC.

- beaucoup trop de codes d'instruction différents.
- taille des instructions élevée et variable (1 à 15 bytes –octet - par instruction)
- structure des instructions non standardisées: exécution complexe, peu performante

Processeur RISC

- Nette séparation entre les instructions d'accès mémoire et les autres
- Instructions standardisées, en taille et en durée d'exécution
- Unité de décodage câblée, non micro codée
- architecture super pipeline, super scalaire
- Très nombreux registres à usage général
- Un ou plusieurs cache (s) internes(s) ainsi que des tampons internes
- et un jeu d'instruction réduit aux instructions simples.

Chaque microprocesseur moderne va chercher des paquets d'instruction X86, les décode sur plusieurs rangées de 8 bits pour les faire passer en instructions RISC. Un problème se pose lors d'instructions conditionnelles. Comment le processeur peut charger plusieurs instructions et les exécuter suivant le choix non prévu au départ. Chaque concepteur utilise soit un circuit de prédiction situé en amont de circuit de décodage CISC - RISC (cas d'INTEL), soit une table de branchement qui mémorise les derniers branchements (AMD). Comme les instructions sont chargées longtemps en avance, il se peut très bien que les données soient modifiées.



Ce changement de type d'instruction a également permis d'implanter une (ou plusieurs) unités de calculs spécialisées en virgule flottante (nombres réels) au sein du processeur. Cette unité de calcul est le processeur mathématique. Généralement, celui d'INTEL est meilleur que celui d'AMD, Cyrix est en dessous de tout. Ce coprocesseur est utilisé dans des calculs en réel (dessin technique et jeux).

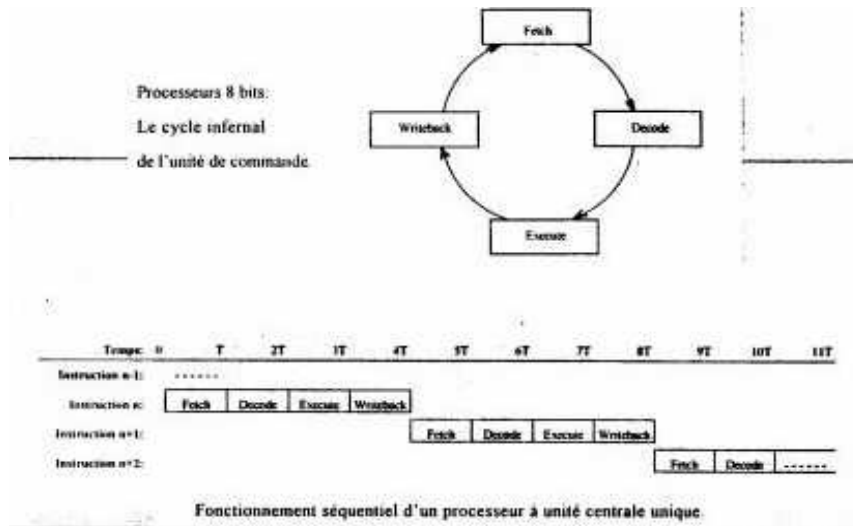
Le passage de processeurs RISC aux processeurs CISC a amené d'autres changements internes, nous allons les passer en revue.

7.5. Architecture pipeline, super scalaire

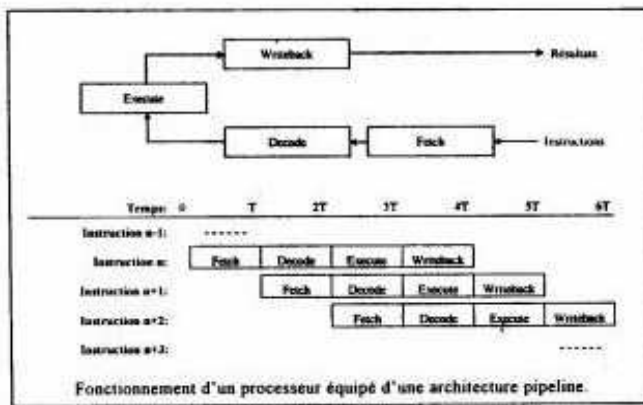
Les 8088-8086 étaient de type non pipeline. Le processeur ne commençait le traitement de l'instruction suivante que lorsqu'il avait terminé la précédente. L'utilisation d'une structure pipeline permet le traitement de plusieurs instructions en même temps. Dans le cas d'une structure pipeline, l'exécution d'une instruction est décomposée en:

1. chargement de l'instruction dans le processeur
2. décodage
3. exécution de l'instruction
4. renvoi du résultat

Comme seules la première et la dernière instruction utilisent les bus externe, le cycle de l'instruction suivante peut commencer en même temps que le décodage interne de l'instruction. Ceci permet d'exécuter plusieurs instructions en même temps.



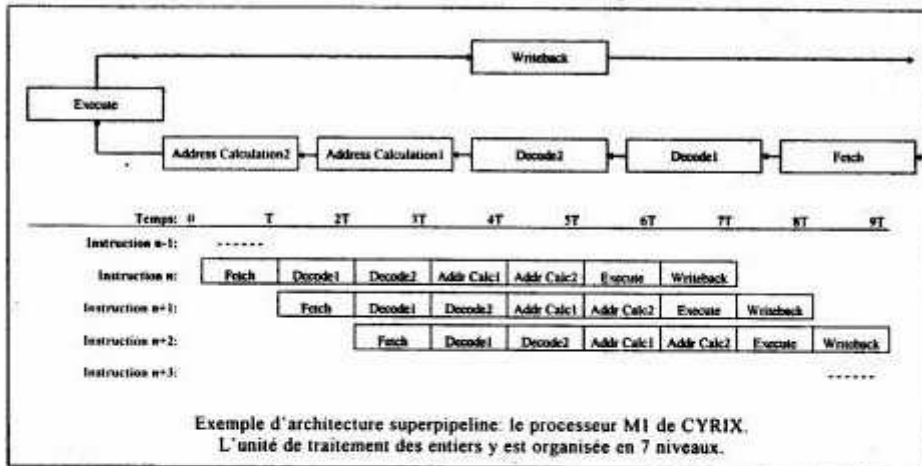
L'architecture pipeline



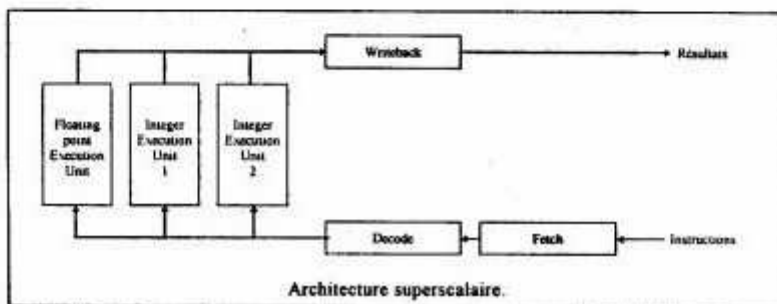
Dans une architecture super pipeline, l'exécution de l'instruction est maintenant divisée par 6. Le nombre d'exécution simultanée n'est guère supérieure en super pipeline par rapport à une architecture pipeline dans les instructions habituelles. Par contre, comme le temps utilisé par chaque partie d'instruction est plus faible, chaque fois qu'une partie est omise (parce que pas nécessaire) l'exécution est encore plus rapide pour la suivante.

Super pipeline: le processeur ne traite pas les instructions d'un seul coup, mais en 6 étapes: pré extraction, décodage d'instruction (2), exécution, génération d'adresse et écriture différée (cas d'un chargement vers ou de la mémoire).

L'architecture superpipeline



L'architecture superscalaire



Dans les structures internes de processeurs super scalaires, le processeur inclue directement plusieurs unités d'exécution d'instruction qui travaillent en même temps. L'Athlon est super pipeline et super scalaire.

7.6. La taille des transistors, alimentation.

Actuellement, les microprocesseurs utilisent des transistors de type CMOS ou BiCMOS. Plus on avance, plus la tension d'alimentation diminue, plus la taille du transistor diminue et plus le nombre de transistors augmente. Une réduction de la taille des transistors diminue la distance de transfert entre 2 informations. Qui dit distance en diminution, dit également moins de dissipations thermiques (puisque le signal suit une distance plus courte, moins d'échauffement dans le processeur). De plus, la complexité des processeurs actuels augmente le nombre de transistors.

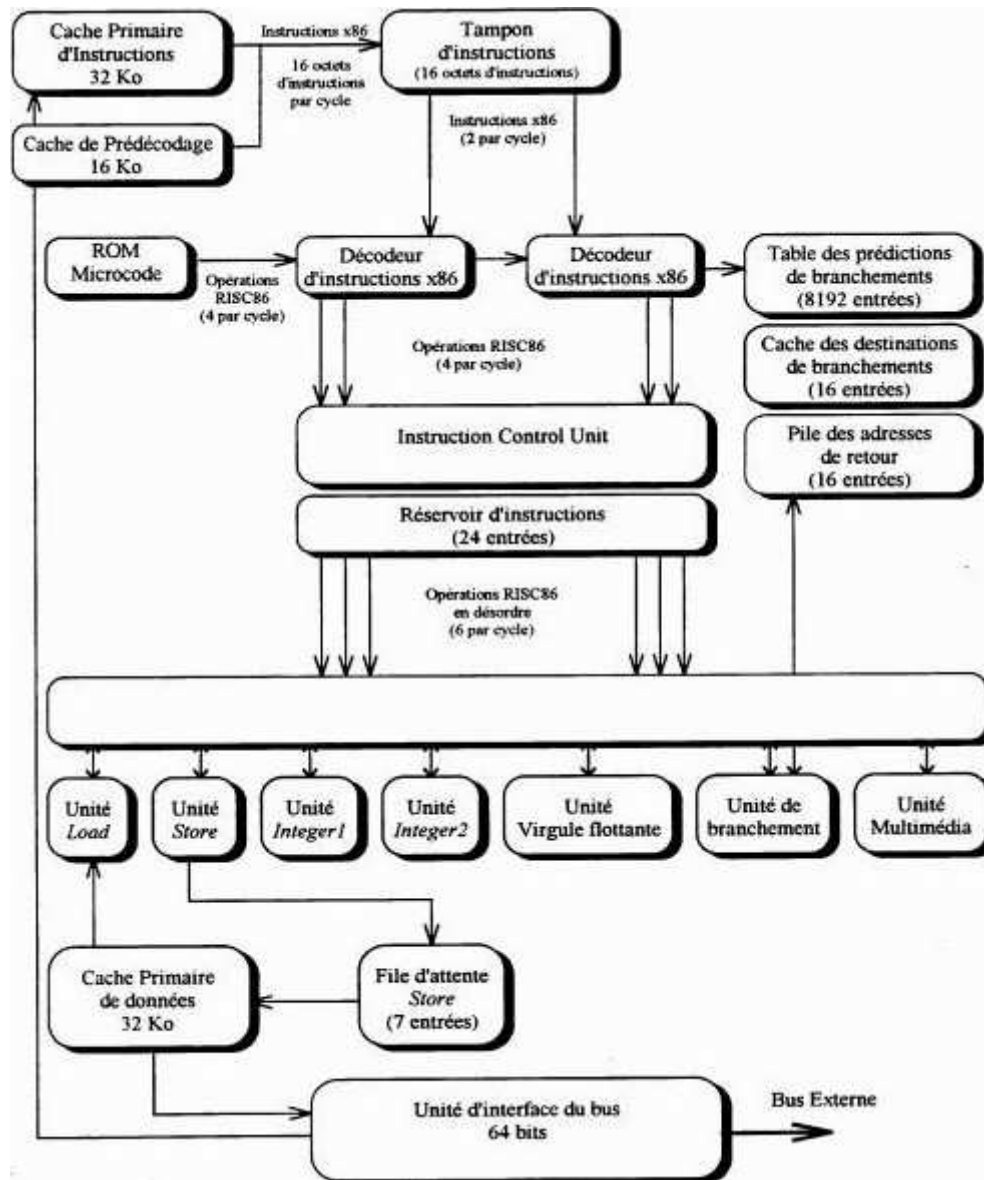
En 1995, la taille d'un transistor était de 0,5 micron. Elle est actuellement de 0,09 micron avec les Pentium IV supérieurs 1,7Gh et derniers athlons.

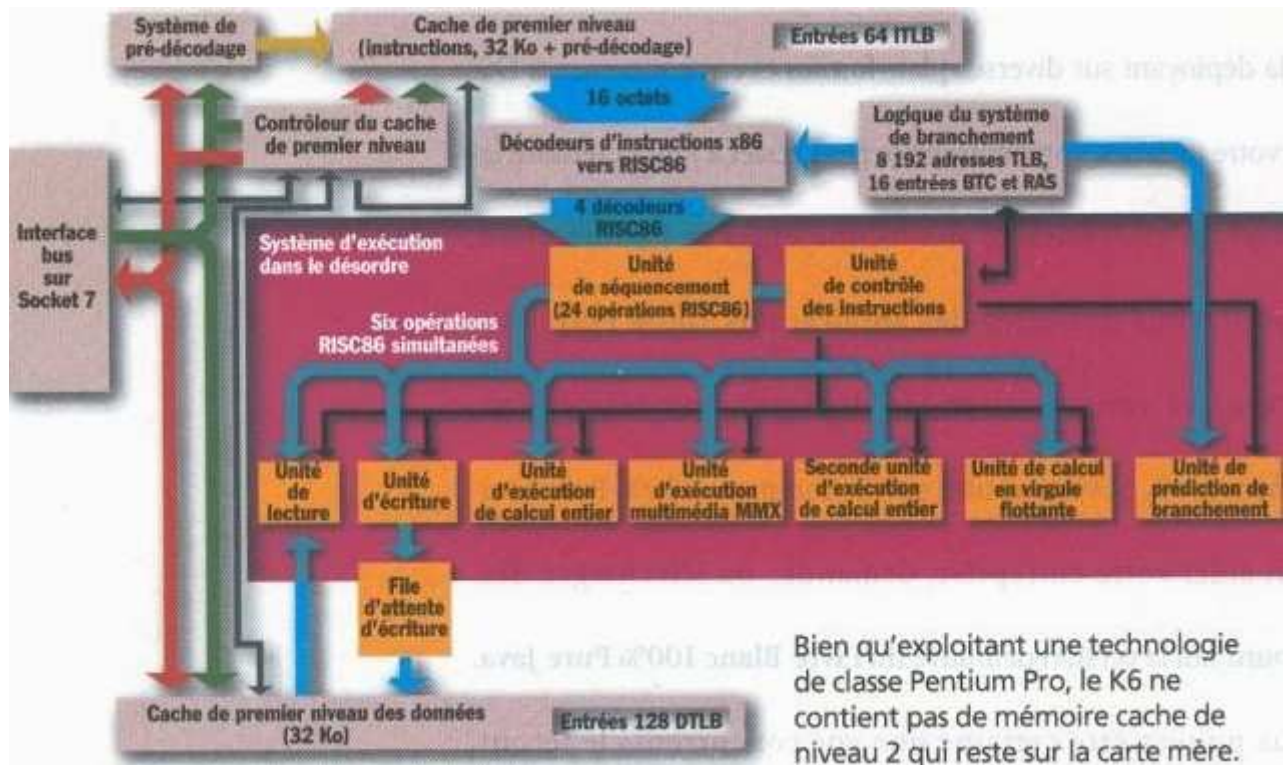
Si depuis les Pentium MMX, les alimentations processeurs sont bi tensions, la tension d'alimentation diminue aussi puisque celle d'un K6-3 à 450 Mhz est actuellement de 2.2 V. Attention, vérifiez systématiquement la tension **sur le processeur** et non sur la carte mère ou sur les notes techniques de celle-ci. Pour les processeurs bi tensions, prendre le V Core, même si la deuxième tension est affichée.

La modification de la tension des processeurs est arrivée avec les premiers 486DX2 de Cyrix (mais est devenue générale). Tous les modèles antérieurs utilisaient le 5V. Une erreur d'alimentation sur un processeur entraîne sa destruction à brève échéance.

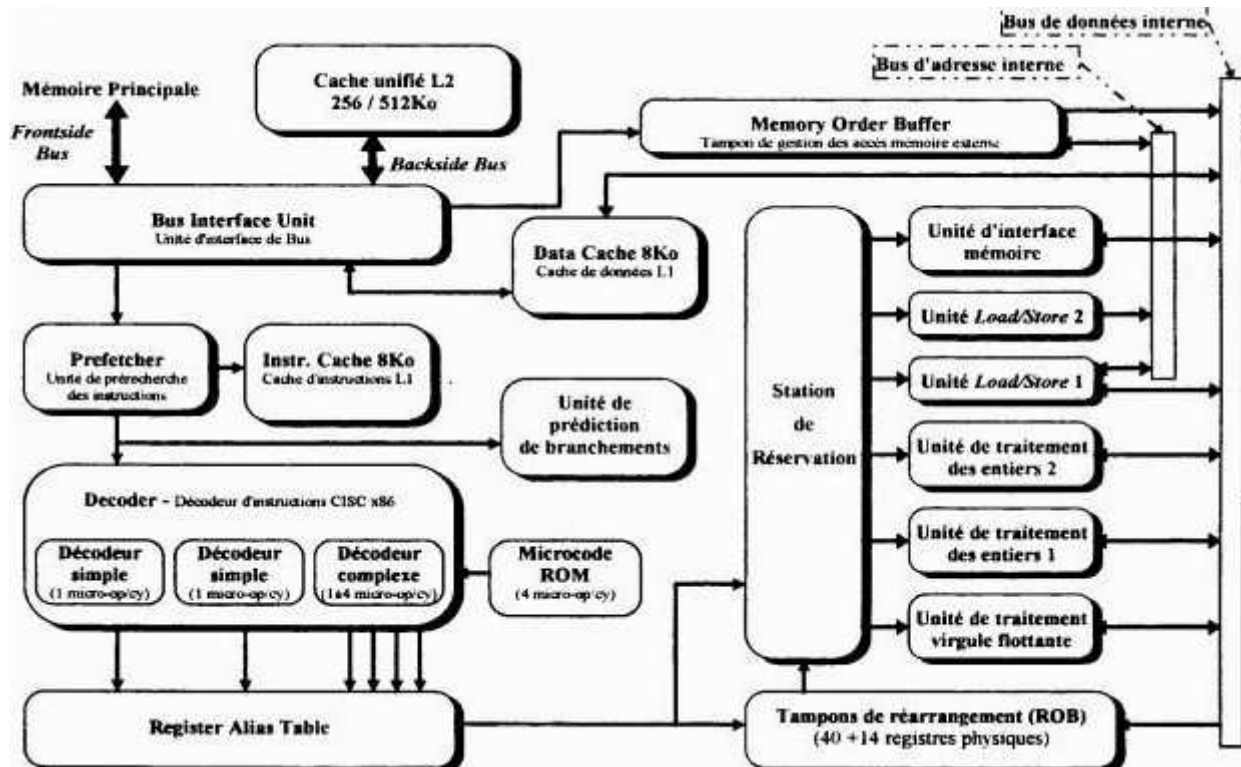
7.7. Architecture interne des processeurs.

En additionnant quelques unes des avancées ci-dessus, on arrive par exemple à un K5 2,5 fois plus rapide qu'un 486 de même fréquence. Analysons brièvement quelques structures internes de processeurs. Commençons par le K6-2 d'AMD ci-dessus.





En haut à gauche, le cache L1 instruction (le cache équivalent de données se trouve en bas) par qui transite toutes les instructions en provenance de la mémoire (cas des AMD). Le système en dessous de pré décodage est utilisé pour les vérifier le type d'instructions qui arrivent (mémoires ou exécution) et sert aux prédictions de branchements. Le tout est relié à un tampon d'instructions qui envoie 2 instructions X86 par cycle d'horloge vers 4 décodeurs d'instructions (passage du CISC au Risc) Ce décodage est exécuté par la ROM Microcode. Les instructions RISC sont transférées vers une table de prédilection de branchements, associé à un autre cache. D'autres part, les instructions sont traitées en RISC par l'instruction Control Unit qui envoie également les instructions à partir d'un réservoir de 24 commandes RISC86 vers les différentes parties en dessous (Load et Store (gestion mémoires), deux unités d'entiers et une de virgule flottante, une unité de branchement qui gère également la table de prédiction et une unité multimédia liée aux instructions MMX). Le tout est relié au bus de donnée externe via le cache de donnée et l'Unité d'interface Bus qui gère également le cache d'instructions. Celle du Pentium Pro ci-dessous est identique, à part le cache L2 (externe) qui n'est pas relié directement à la mémoire. Cette structure est reprise dans les Pentium II et Pentium III.



Structure d'un microprocesseur Pentium Pro

7.8. En conclusion

La modification de la structure interne des processeurs rend les performances d'un PC assez difficile à vérifier en se basant uniquement sur la fréquence du processeur. Ceci explique les bricolages commerciaux d'AMD qui désigne les vitesses de ses microprocesseurs, non pas par la vitesse effective, mais par des vitesses de comparaison par rapport aux équivalents INTEL en fonction des Benchmarks. Cyrix avec les 586 faisait de même. Si le nombre d'instructions décodées en même temps devrait encore augmenter, la direction actuelle des processeurs va vers la diminution de la taille des transistors. Ceci permet d'augmenter la fréquence des processeurs. En effet, plus le nombre de transistors augmente, plus la dissipation de chaleur est importante. Mais plus la taille d'un transistor diminue, moins cette dissipation est importante. D'autres matériaux sont en cours d'utilisation, IBM utilise déjà des conducteurs internes en cuivre, alors que la technologie actuelle est basée sur l'aluminium. Le cuivre est meilleur conducteur électrique (d'où moins de dissipations thermiques), mais nécessite une technologie plus pointue, avec des difficultés d'intégration plus importante.