Chapitre 3 :

Les machines parallèles

1- Introduction

Les progrès technologiques se sont accélérés de façon significative depuis plus d'une quarantaine d'année (milieu 70).

En très peu de temps, on est passé de la 3è génération à la **5ème**, voire la **10ème** génération.

Les composants de la machine ont été très réduits en taille, et leur puissance a été de plus en plus accélérée, ce qui a conduit les concepteurs à imaginer des architectures très variées, et assez loin de la machine traditionnelle et classique, à la Von Neumann

En particulier, les intérêts se sont portés, après les VLSI de la 4ème génération (et après les transistors, circuits imprimés, circuits intégrés) sur le fait de faire coopérer plusieurs machines dans une architecture : c'est le domaine des *machines parallèles*.

Si l'on adresse un bref panorama de divers types d'architecture parallèle, on trouve :

1- les architectures les plus simples, obtenues par assemblage de processeurs du commerce : ce sont les multiprocesseurs, ou réseaux

2- les machines destinées au calcul scientifique, qui permettent un traitement toujours plus performant des calculs vectoriels et matriciels

3- les machines (les plus récentes), aux possibilités encore inexplorées : les machines à base de transputers et « la connexion machine »

Grâce à ces machines, il est possible d'obtenir une exécution concurrente de plusieurs programmes en reliant plusieurs processeurs.

Le fonctionnement autonome des processeurs permet l'exécution simultanée de plusieurs programmes séquentiels.

Un outils de connexion approprié permet d'effectuer les synchronisations et communications nécessaires.

Cette description générale répond à de nombreuses machines. Mais celles-ci se distinguent par les processeurs composant ces architectures, et surtout suivant les moyens de communication inter processeurs.

C'est ce dernier critère qui nous permet de définir 2 grandes classes :

1- les **multiprocesseurs** : les communications s'effectuent via une mémoire partagée

2- les **architectures réparties** : les communications s'effectuent exclusivement par des réseaux mis en oeuvre par des câbles, des fibres optiques.

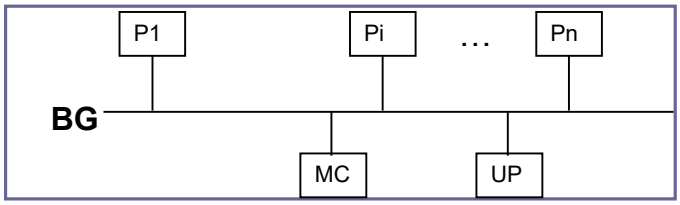
Ces architectures sont de 2 types, avec des performances différentes :

- **Réseaux locaux** : ne permettent de relier que des processeurs voisins, distants de quelques centaines de mètres.

- **Réseaux « généraux »** : permettent de relier des processeurs situés à des distances quelconques.

2- les multiprocesseurs

Sont représentés par la figure suivante :



Où **Pi** : processeurs homogènes
MC : mémoire commune à laquelle accèdent tous les processeurs Pi
BG : Bus Global
UP: Unités Périphériques

Les unités périphériques sont accessibles via BG

Cette architecture, assez simple, peut être gérée par un système d'exploitation qui ne diffère pas beaucoup d'un système de multiprogrammation pour mono-processeur (concurrent).

La tâche principale est le partage des ressources qui sont

- le bus,
- la mémoire,
- et les unités périphériques.

La difficulté majeure dans ce cas est que plusieurs calculs peuvent s'exécuter simultanément sur les divers processeurs, et peuvent effectuer des requêtes d'accès à la mémoire commune, en vue de lecture et surtout de mise à jour.

Exemple : de nombreuses machines multiprocesseurs ont été développées (séquent, Alliant, Encore, Inter-technique), avec l'utilisation des nouveaux processeurs tels que Motorola 680xx, NS 32000, …

Exemple : pour les logiciels, les systèmes « opératoires » proposés offrent toujours une interface compatible avec le système le plus courant actuellement (UNIX).

Deux (2) approches existent :

1- construction d'un Unix pour machine multiprocesseurs, c'est-à-dire, mise en œuvre de façon à tirer profit des possibilités de parallélisme (Unix de Encore Computer, Dynix de Séquent, …)

2- construction d'un système exploitant au mieux les possibilités de parallélisme, et offrant une interface UNIX.

Exemple : CHORUS très différent de Unix conceptuellement, mais peut être utilisé grâce aux commandes Unix

3- Les architectures informatiques réparties

Ces dernières décennies, l'emploi des micro-ordinateurs s'est généralisé, car ces derniers offrent une autonomie de calcul avec des ressources propres (mémoires et périphériques).
La connexion des micros est encore plus intéressante, pour les raisons qui suivent :

- communication entre plusieurs usagés.
- capacité mémoire des micros (assez réduite en général).
- mise en œuvre difficile des grosses applications, demandant temps et mémoires.

C'est ainsi que l'idée des systèmes informatiques répartis est née.

Cette idée repose sur un réseau qui permet
d'interconnecter
des postes de travail individuels
et des serveurs spécialisés
grâce
à des outils de communication appropriés.

Les **caractéristiques** des architectures réparties sont principalement :

a. le degré de parallélisme élevé

b. absence de mémoire commune (impossibilité de définir un état global du système, comme on peut le faire dans les systèmes monoprocesseurs et multiprocesseurs)

C problèmes de fiabilité accrus : Pb de résistance aux pannes, dus à la multiplicité de ressources présentes dans le système et la mauvaise fiabilité relative des moyens de communication

Il existe trois (03) **types d'architecture**, en fonction du degré de couplage entre les divers processeurs :

- **fort couplage** (machines multiprocesseurs sans mémoire commune)
- **faible couplage** (réseaux locaux)
- **réseaux grande distance**

a) Architectures à fort couplage : sont constituées d'un grand nombre de processeurs reliés entre eux suivant une topologie particulière.

Exemple : IPSC : machine dans laquelle les processeurs (INTEL 80286) sont connectés selon une topologie « Hypercube » : Un hypercube est un cube de dimension n comportant 2n nœuds, où chaque nœud est relié à n voisins.

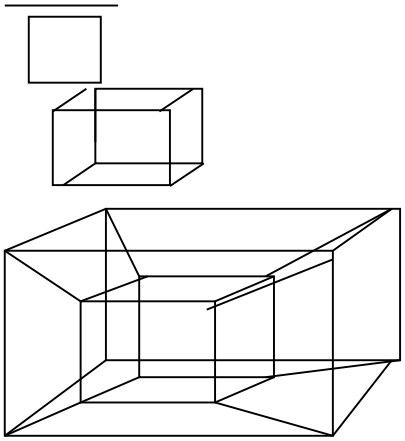
Dim 0 1 nœud.

Dim 1 2 nœuds

Dim 2 4 nœuds

Dim 3 23 = 8 nœuds

Dim 4 24 = 16 nœuds



Avantage de cette topologie :

- extensible : un cube de dimension n peut être formé grâce à 2 cubes de dimension n-1, où les nœuds ont été reliés 2 à 2

- la distance moyenne entre 2 nœuds n'est que de n/2 et la distance maximale est n, où n est la dimension du cube. Chaque nœud peut donc y être accédé en un petit nombre de transmissions de message.

- de nombreuses architectures (chaînes, anneaux, arbres,…) peuvent être décrites sur cette topologie.

b) Réseaux locaux : Possèdent trois caractéristiques principales :

- couvrent une faible étendue géographique (moins d'un kilomètre en moyenne)

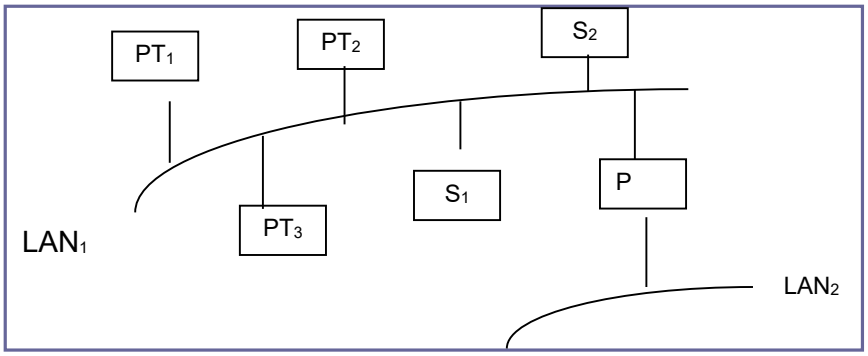
- leur débit est élevé (10 à plusieurs centaines de Méga bits)

- sont bien intégrés, c à d, conçus comme un tout.

Exemples :

- **ETERNET** (développé initialement au centre de recherche de XEROX à Paolo Alto).
Les messages sont diffusés à tous les hôtes qui sont en permanence prêts à les recevoir.
Des problèmes surviennent du fait que plusieurs émetteurs peuvent envoyer un message simultanément, auquel cas, des phénomènes de brouillage se produisent.

- Anneaux à jeton : un jeton (message spécial) circule sur un réseau bouclé (en anneau).
Seul le site possédant le jeton peut émettre un message.
Un site recevant le jetant alors qu'il n'a rien à émettre le transmet au site suivant. (Exemple de tel réseau : le **Cambridge Ring**)



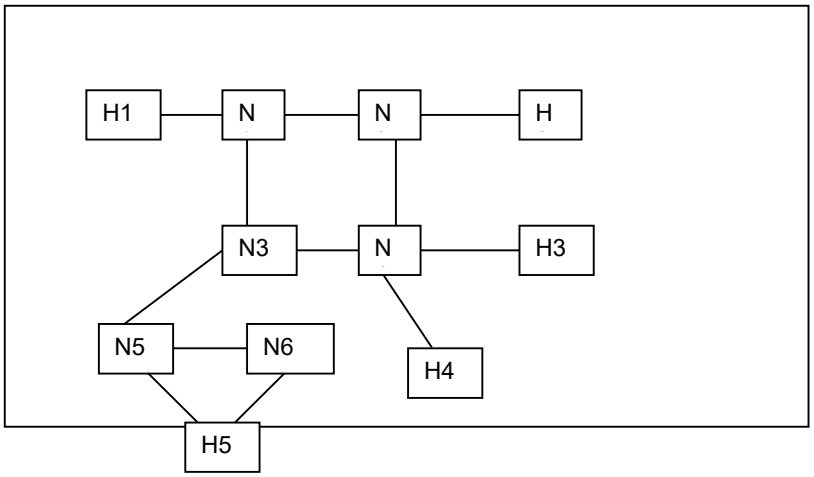
PTi : les postes de travail
Si : serveurs (exemple : serveurs d'impression ou serveurs de fichiers)
P : une passerelle permettant de relier 2 réseaux locaux (LAN1 et LAN2)

c) Réseaux grande distance : est une collection :

- de calculateurs hôtes
- et de nœuds de communication

} pouvant être très éloignés les uns des autres (entre 20 et 10.000 km).

Les liaisons entre les divers composants du réseau peuvent s'effectuer par:
- câble,
- par liaison radio
- et par satellite



Hi : ordinateurs de calcul
Ni : nœuds de communication

Il existe 2 types de réseaux grande distance :

a- réseaux à commutation de circuits :

Un lien physique est établi dans chacun des nœuds entre voie d'entrée et voie de sortie.
Ce lien est maintenu durant toute la communication.

b- Réseaux à commutation par paquets :

où l'unité de transmission est le paquet de taille fixe.
Les messages sont découpés en paquets lors de l'émission, puis reconstruits lors de la réception.
Ceci améliore les performances des systèmes de transmission : un paquet peut transiter par un nœud indépendamment des autres paquets constituant le message logique. En cas d'erreur de transmission, seul le paquet fautif doit être retransmis.

4 - Les machines massivement parallèle

Ce sont des supercalculateurs dont la puissance de calcul est phénoménale, dont la rapidité va du Mflop (million d'instructions flottantes par seconde) au Gflop (milliard…)

La course à de telles machines n'est pas encore terminée, car les besoins en puissance de calcul ne semblent pas avoir de limites. Ces supercalculateurs permettent de traiter des applications jusqu'alors irréalistes parce que nécessitant des temps de calcul prohibitifs (calcul numérique, génétique, …)

Par exemple :
L'expérimentation des phénomènes en laboratoire est impossible parce qu'elle porte sur des durées trop courtes (domaine nucléaire) ou trop longues.
Une bonne modélisation des phénomènes et leur simulation deviennent alors un excellent substitut.
Mais le traitement informatique de tels modèle demande un temps de calcul considérable, d'où la nécessiter de posséder des machines très puissantes (domaine génétique).

Ces supercalculateurs, soit en usage, soit en cours de développement, sont de 2 classes principales : et on trouve aussi les MIMD

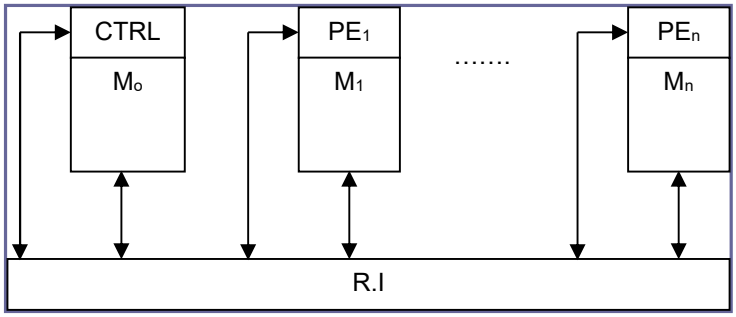
- les calculateurs tableaux
- les calculateurs vectoriels
- les calculateurs à flots d'instructions et à flots de données.

a) Les machines tableaux :

Sont appelés **SIMD** (Single Instruction Multiple Data) signifie qu'une même instruction est exécutée simultanément par tous les processeurs sur des données différentes.
Ceci implique la nécessité de plusieurs unités arithmétiques et logiques

Son organisation générale est schématisée comme suit :

PE i (i=1, …, n) : unités de traitement, avec leur mémoire
Mi : mémoire de l'unité PE i
RI : réseau d'interconnexion, relie les diverses unités



CTRL unité de contrôle permettant de contrôler l'ensemble de l'architecture.
M0 sa mémoire
Les PEi sont identiques et exécutent de façon synchrone la même instruction, sur des données situées dans leur propre mémoire.

Le programme exécuté par ces unités de traitement est situé dans la mémoire Mo du processeur de contrôle, qui distribue de façon synchrone aux PEi, les instructions à exécuter.

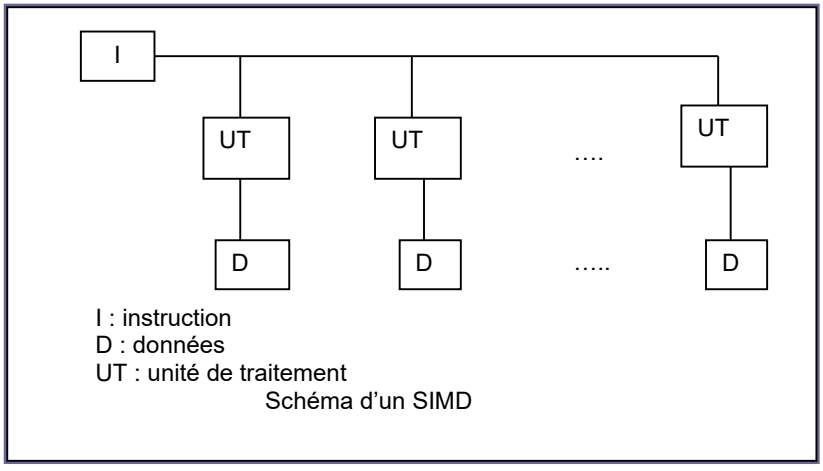


Schéma d'un SIMD

Ces calculateurs sont adaptés à une classe de problèmes très spécifiques qui nécessitent des calculs fréquents sur des tableaux (ou matrices) de données.

Pour permettre une exploitation optimale d'une machine tableau, il est nécessaire que le programmeur prête une attention particulière à la structure de ses algorithmes et à l'organisation de ses données.

Exemple de machines tableaux

MPP (Massively Parallel Processor) au début des années 80 :

- comporte 16384 unités de traitement (tableau 128 x 128 éléments)
- et peut effectuer 430 millions d'additions flottantes (32 bits) par seconde et 216 millions de multiplications flottantes (32 bits) par seconde.

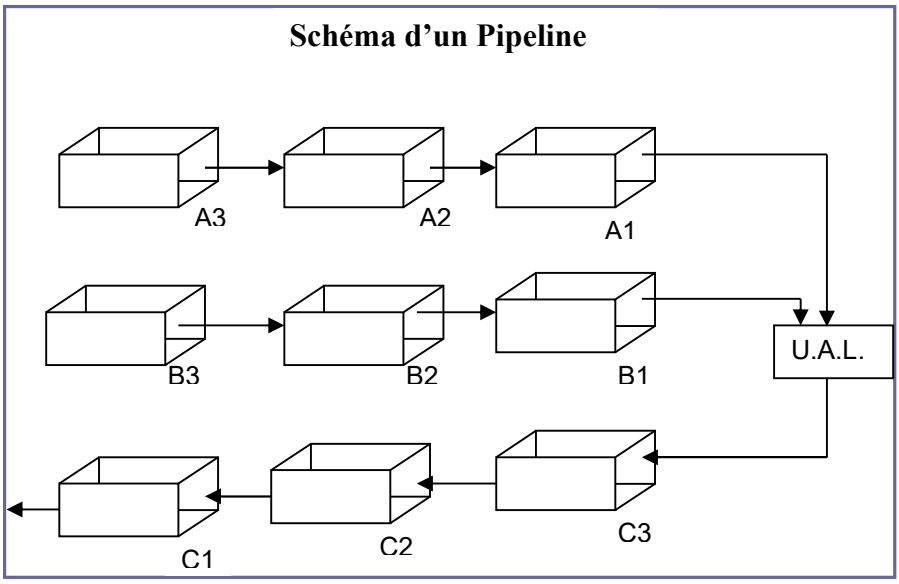
Application :
- Traitement images satellites
- Prévisions météorologiques
- Simulateurs aérodynamique
- Traitement signaux radars
- Traitement images en général.

b) les calculateurs vectoriels :

Leur caractéristique essentielle est d'être très efficaces sur les opérandes de type vectoriel, grâce à l'utilisation d'une technique de « pipe-line », technique qui permet

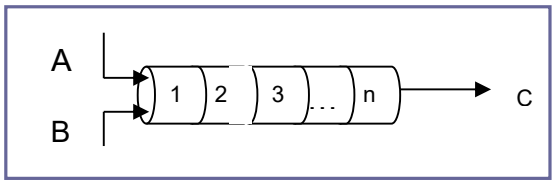
- la recherche,
- le décodage,
- et l'exécution de plusieurs instructions simultanément.

Le matériel étant divisé en étages, plusieurs instructions peuvent être simultanément en cours d'exécution à des étages différentes.



Les Ai, Bi et Ci sont dans des registres de travail.
L'U.A.L est unique mais découpée (en blocs de 12,5 ns) et multiplexée.

Le fonctionnement : en même temps que l'opération des deux opérandes A1 et B1 se fait, les autres opérandes se décalent dans les registres suivants. Après un certain temps d'amorce un premier résultat est obtenu puis d'autres sont obtenus par période d'horloge (tp = 12,5 ns)



Le nombre maximum de ces étages borne le parallélisme exploitable.
Pratiquement tous les supercalculateurs sont structurés en pipeline.
La puissance du CRAY 1 est de 25 à 100 M flops avec un parallélisme synchrone.

Exemple de tel calculateur : **CRAY** 1.

Cette machine est composée de 12 unités fonctionnelles indépendantes, donc capables d'effectuer des calculs en parallèle.
Ces unités fonctionnelles sont réparties en 4 groupes :

- unités traitant des vecteurs
- unités arithmétiques flottantes
- unités arithmétiques scalaires
- unités de traitement d'adresses.

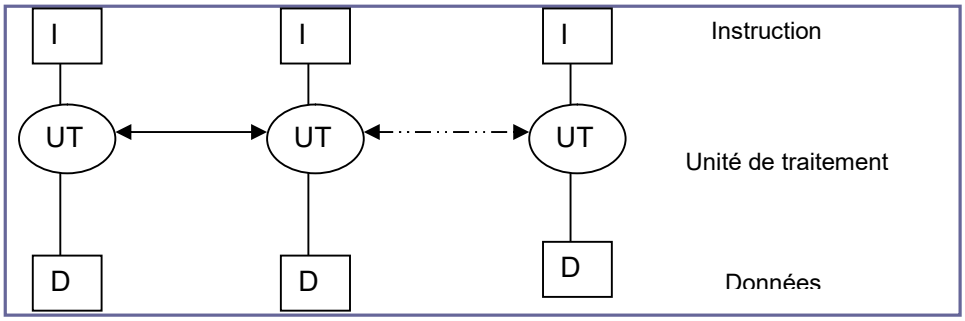
La mémoire principale de CRAY 1 est constituée de 16 bancs indépendantes et a une capacité totale de 1 Mmots de 64 bits, il existe 3 tampons d'instructions, chacun possédant une capacité de 64 instructions de 16 bits.
Chaque tampon est relié à la mémoire principale via un bus de données d'une largeur de 64 bits.
Il existe 3 ensembles de registres :

8 registres d'adresses (24 bits),
8 registres scalaires virgule flottantes (64 bits)
et 8 registres vectoriels pouvant contenir chacun 64 valeurs flottantes représentées sur 64 bits.

Les performances du CRAY 1 sont de l'ordre de 100 Mflops.
Les E/S sont effectuées via 24 canaux (12 pour les entrées, 12 pour les sorties).
Le taux maximal de transfert est de 10 Mmots par seconde et par canal.

c) MIMD (Multiple Instruction – Multiple Data)

Le matériel



Entre les différentes unités de traitement, il existe une coopération.
Le fonctionnement : caractérisé par la coopération de processeurs qui

- exécutent des programmes distincts,
- se partagent une mémoire
- et se synchronisent de temps en temps pour échanger des résultats.

En général, il existe le problème de synchronisation de plusieurs événements.

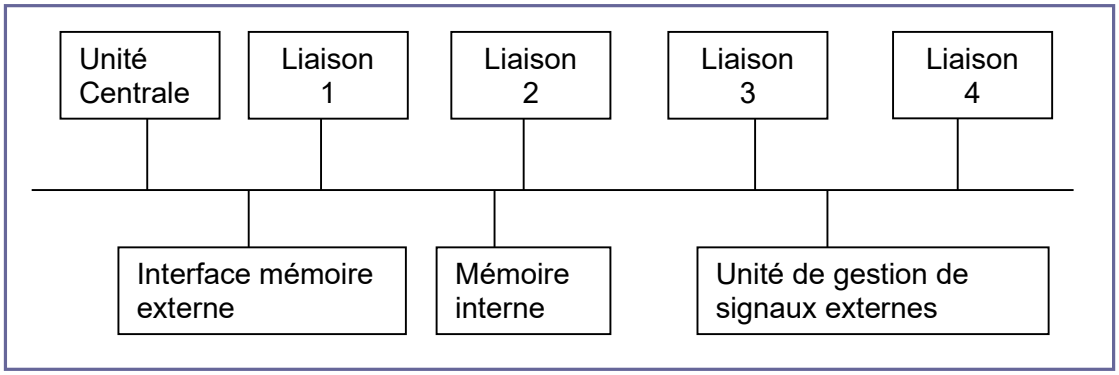
5. Architectures prometteuses

a) Machines à base de transputers

Le transputer est une machine originale développée par la firme Britanique IMNOS.
Cette machine a l'originalité d'être spécialement conçue pour exécuter efficacement un langage parallèle de haut niveau :

Le langage **OCCAM**, c'est donc une machine langage OCCAM.

Structure du transputer :



Le premier transputer, le T414 possède

- une unité centrale de 32 bits,
- une mémoire interne de 2 k octets,
- une interface mémoire externe 32 bits
- et 4 liaisons permettant notamment des connexions avec d'autres transputers. Les liaisons ont un débit de 20 M bits.

Le **T414** permet d'obtenir des performances de l'ordre de 10 millions d'instructions par seconde sur des programmes séquentiels.

Comme indiqué, chaque transputer est doté de 4 liaisons à partir desquelles il est possible de construire un réseau de transputers - c'est là une originalité des transputers : permettre de construire une architecture sophistiquée par assemblage de composants de base.

Dans le Projet Européen ESPRIT, des Britanniques et Français ont étudié une telle machine parallèle : Supernode, composée de 16 transputers.

Les supernodes peuvent à leur tour être connectés entre eux et constituer ainsi des réseaux à la demande (on peut aller jusqu'à 64 supernodes = 1024 transputers).

Applications développées sur le supernode, avec le langage OCCAM :

- Traitement d'images
- CAO de circuit VLSI
- Traitement du signal

Les performances ont été encourageantes

b) La connexion machine

Dans ce qui a été présenté, les machines peuvent comporter quelques centaines de processeurs.
La connexion machine, comporte **65536 processeurs** pouvant coopérer à un même calcul.
Ces 65536 processeurs sont regroupés en 4096 groupes de 16 processeurs.
Ces **groupes** sont reliés suivant le modèle de l'hypercube de dimension 12.
A chaque processeur est associé une mémoire locale de seulement 4096 bits.

L'idée consiste à associer un processeur à chaque élément de donnée, chaque cellule de mémoire étant ainsi capable de résoudre un calcul local grâce à ce processeur. Un canal d'E/S à très haut débit (500 Mbits/s) peut transférer les données directement de la connexion machine vers différents périphériques.
La connexion machine est de type SIMD.
Les performances théoriques maximales de cette machine sont de 2500 M flops.

Application :

- Dynamique des fluides
- Analyse des données
- Reconnaissances des formes

Conclusion :

On est encore loin de maîtriser la construction de programmes pour de telles machines, mais une nouvelle algorithmique est en train d'apparaître, permettant de résoudre de façon très efficace des problèmes classiques.
Il existe aussi des

- mini-supercalculateurs,
- des machines systoliques,
- des machines neuronales.