



Institut National des Sciences Appliquées
Ecole publique d'ingénieurs

Cours Parallélisme

I. Parallélismes

Jean-Louis Pazat



Plan

2

- Introduction
- Parallélisme d'expression
- Applications
- Parallélisme d'exécution
 - Architecture de machines
 - Influence sur l'expression du parallélisme
- Challenges

- Définition générique
 - Activité parallèle

activité qui existe, s'exerce
en même temps qu'une autre

- Activité ?
- En même temps ?

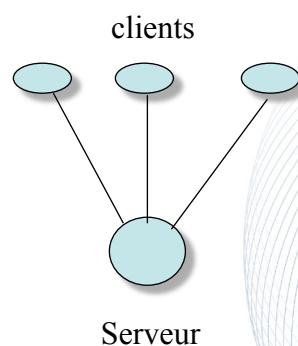
- Parallélisme d 'exécution
 - machines « parallèles »:
 - multiples unités fonctionnelles
 - Machines en réseau, machines multiprocesseurs
 - DMA, processeurs spécialisés
 - à l 'intérieur d'un processeur
- Parallélisme d 'expression
 - organisation « naturelle » d 'une application
 - découpage d'un algorithme

- Introduction
- Parallélisme d'expression
 - Applications
- Parallélisme d'exécution
 - Architecture de machines
 - Influence sur l'expression du parallélisme
- Difficultés

- Dû à la distribution « physique »
- Dû à la nature du problème
- Dû à la conception
- Dû à l 'algorithme
- Pour l 'amélioration des performances

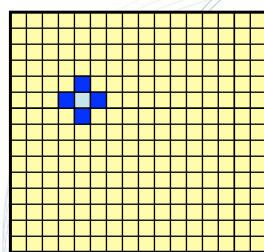
- client / serveur géographiquement répartis
 - WWW, agents, (web) services
- Automatismes répartis
 - automobile, appareils photos, ...
- Surveillance
 - militaire, géologique, météo, marine, ...
- Distribution électronique de journaux

2 grandes catégories



P2P : totalement distribué

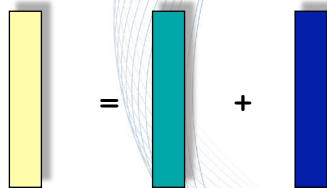
- Dû au modèle
 - simulation
 - calculs en physique/chimie
 - interactions entre particules (relaxation)
 - simulation de foules
 - automates cellulaires
 - agents (IA)



$$A(i,j) = f(A(i-1,j), A(i+1,j), \\ A(i, j+1), A(i, j-1), \\ A(i,j))$$

- Expression de calcul indépendants

pour tout i dans $[1..N]$ faire
 $A[i] = B[i] + C[i]$



- Diviser pour résoudre

$$P = \{p_i\}, i [1, N]$$

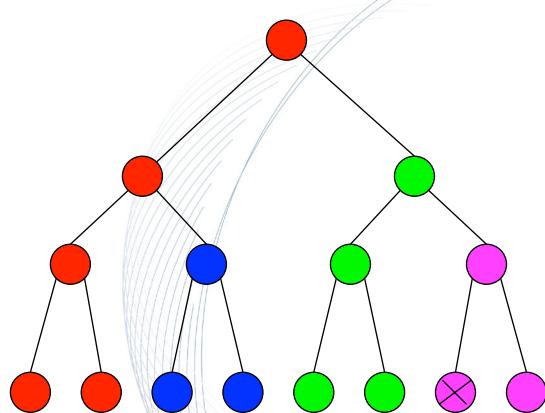
$$\begin{aligned} & \{p_i\} Ci \{q_i\}, i [1, N] \\ & \{\Lambda q_i\}(C_1, C_2, \dots, C_N)\{\Lambda q_i\} \end{aligned}$$

- p_i indépendants \rightarrow idéal
- prendre en compte les dépendances

Exemple de parcours d'arbre

13

- P0
- P1
- P2
- P3

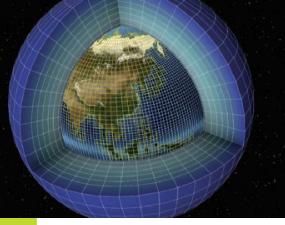


- Exemples : PSEP (Branch and Bound)

Pour l'amélioration des performances

14

- Performances recherchées
 - diminution le temps de calcul
 - augmentation de la taille des problèmes traités
- Paralléliser un calcul
 - pour son exécution sur
 - une machine parallèle spécialisée ou non
 - une « grappe de PC »
 - un réseau à très grande échelle



Ordre de grandeur : exemple simulation du climat 15

Source : CEA Saclay

- Temps de calcul
 - une simulation type sur la période 1750-2100 utilise huit processeurs NEC qui tournent pendant deux mois.
 - Calcul sur 150 millions de pas de temps
 - On peut par ex. calculer la force et la direction du vent toutes les minutes sur 350 ans.
- Dimension des problèmes
 - 180 x 150 points à la surface x 19 hauteurs dans l'atmosphère (513 000 cellules).
 - Les chercheurs espèrent atteindre une définition de 360 x 200 x 40 points dans les années à venir (2 880 000 cellules).
 - Au final, les simulations engendrent 100 téraoctets de données



• Introduction • Parallélisme d'expression • Applications • Parallélisme d'exécution - Architecture de machines - Influence sur l'expression du parallélisme • Difficultés 16

Institut National des Sciences Appliquées
Ecole publique d'ingénieurs

- Sciences de la vie
- La terre, l'environnement, la société
- Evolution de l'univers
- Physique, chimie
- Nouvelles technologies, processus industriels
 - > vidéos guitare + Hurricane

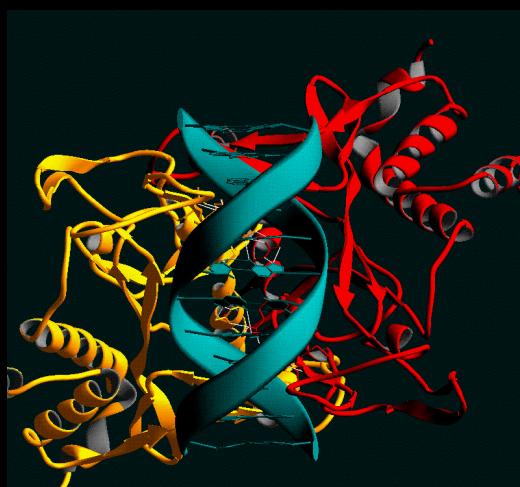
Protein-DNA recognition

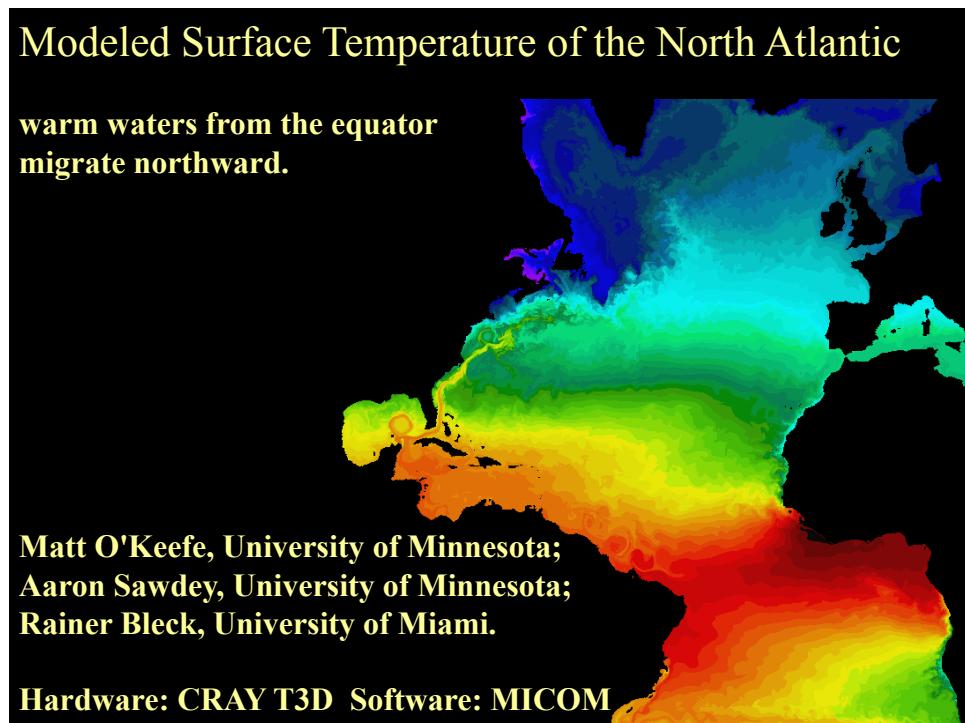
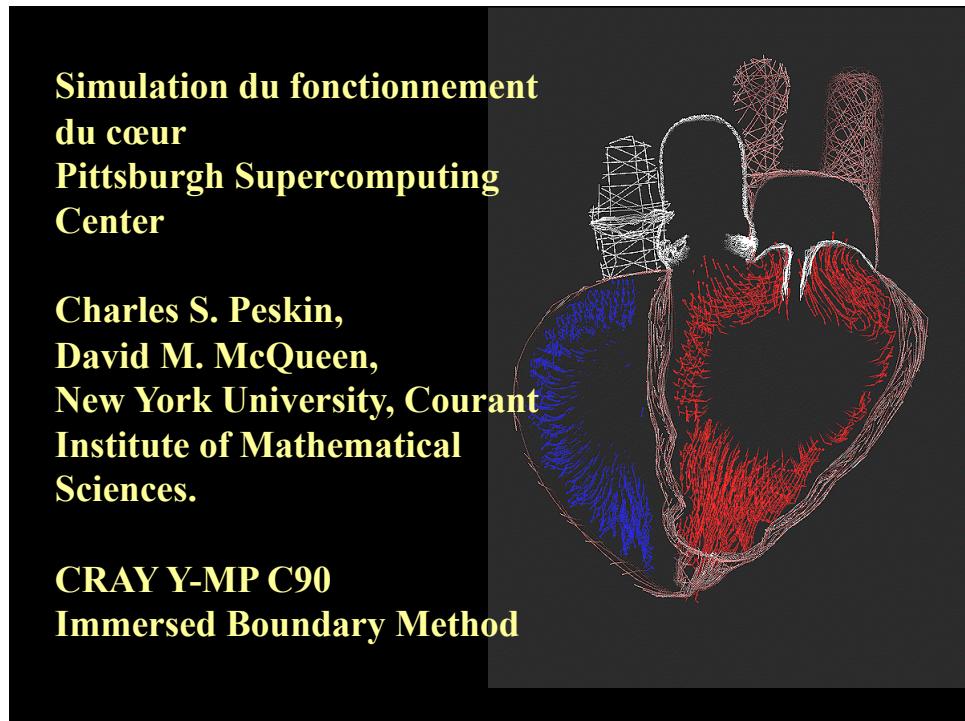
Detailed model of the structure of Eco RI.

John M. Rosenberg,
University of
Pittsburgh.

CRAY Y-MP

X-PLOR, AMBER





- Introduction
- Parallélisme d'expression
- Applications
- **Parallélisme d'exécution**
 - Architecture de machines
 - Influence sur l'expression du parallélisme
- Difficultés

- Quelques exemples de machines

Earth Simulator

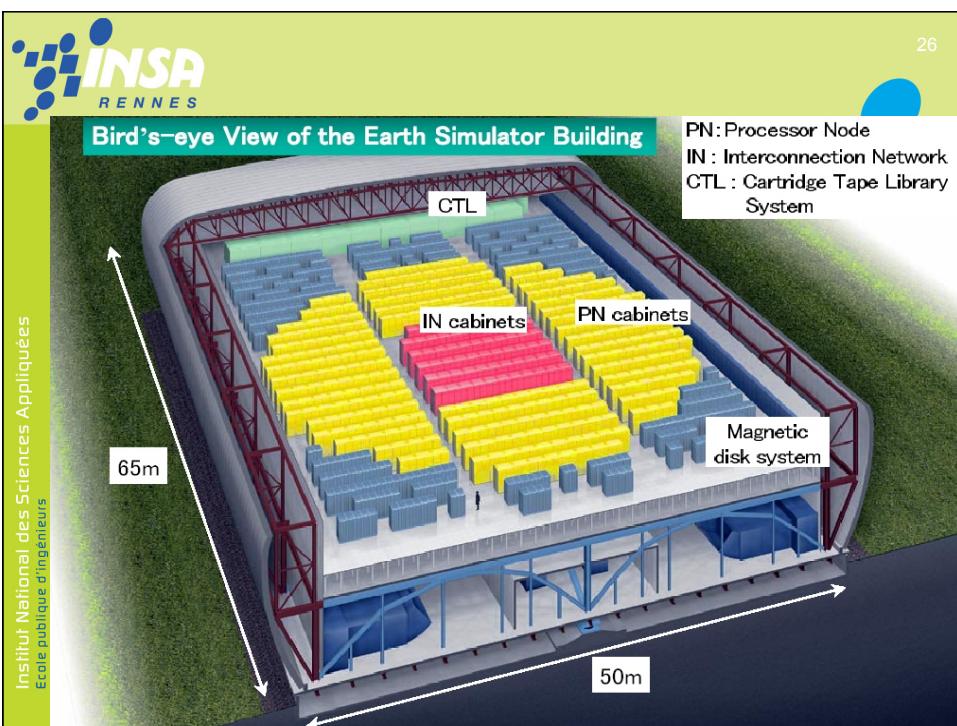
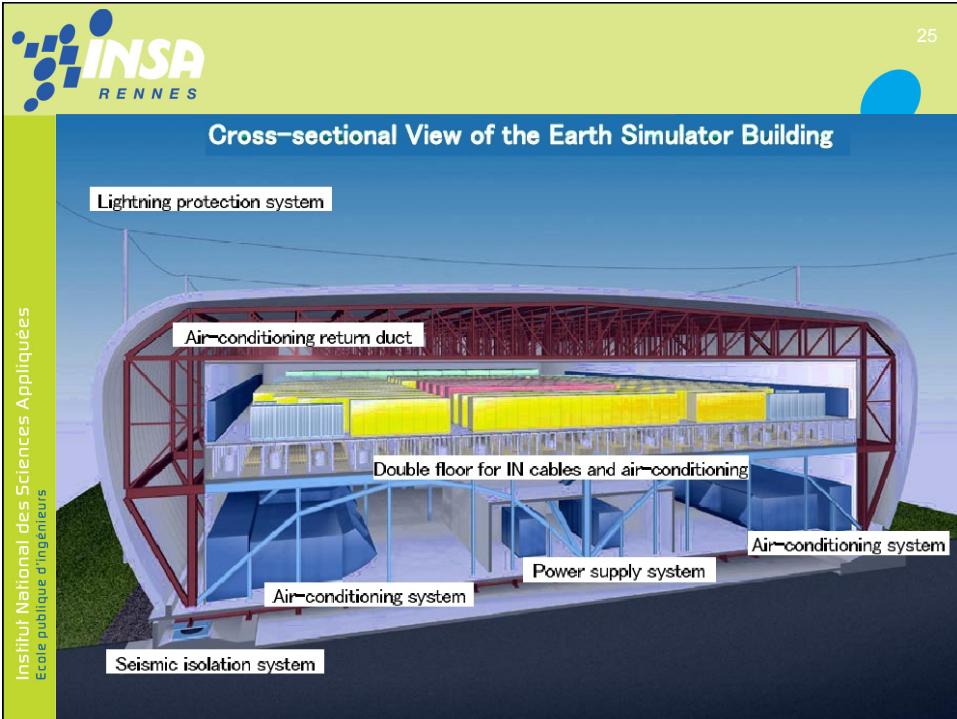
23

- 40 Tflops, 10 Toctet
- Dimensions physiques
 - 50m x 65 m
 - 320 armoires (nœuds)
 - 65 armoires : réseau
- 640 nœuds
 - 8 processeurs vectoriels / nœud
 - 8 Gflops par processeur
 - 64 GFlops
- 5120 processeurs

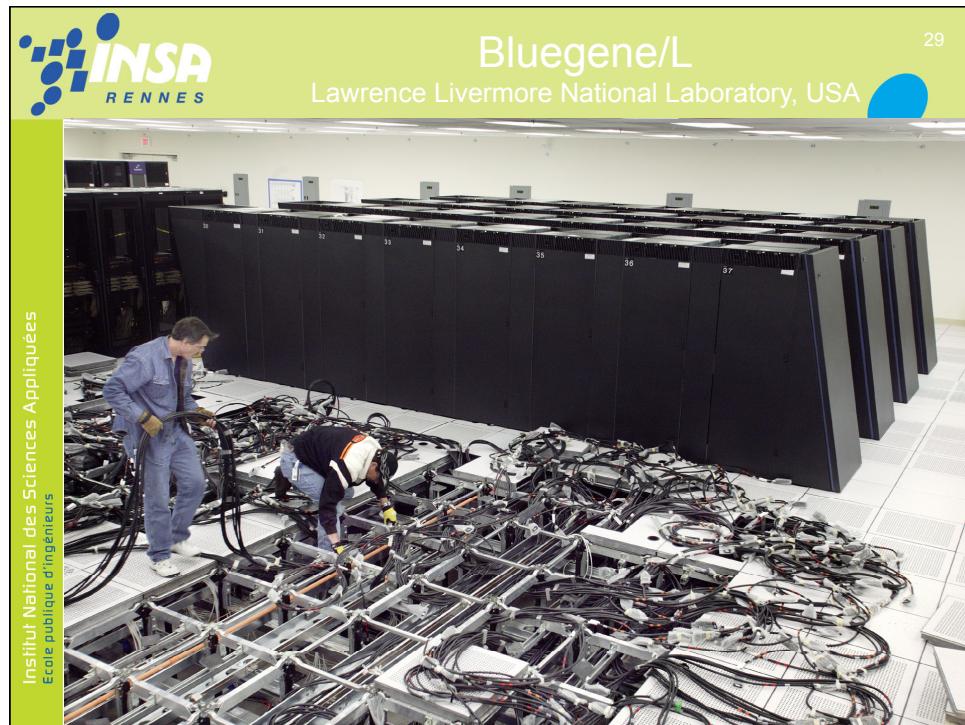
Conceptual drawing of the Earth Simulator facilities at completion

24









TOP 500 Novembre 2008

31

Rank	Site	Computer
1	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband IBM
2	Oak Ridge National Laboratory United States	Jaguar - Cray XT5 QC 2.3 GHz Cray Inc.
3	NASA/Ames Research Center/NAS United States	Pleiades - SGI Altix ICE 8200EX, Xeon QC 3.0/2.66 GHz SGI
4	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution IBM
5	Argonne National Laboratory United States	Blue Gene/P Solution IBM
6	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron QC 2.3 Ghz, Infiniband Sun Microsystems
7	NERSC/LBNL United States	Franklin - Cray XT4 QuadCore 2.3 GHz Cray Inc.
8	Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz Cray Inc.
9	NNSA/Sandia National Laboratories United States	Red Storm - Sandia/Cray Red Storm, XT3/4, 2.4/2.2 GHz dual/quad core Cray Inc.
10	Shanghai Supercomputer Center China	Dawning 5000A - Dawning 5000A, QC Opteron 1.9 Ghz, Infiniband, Windows HPC 2008 Dawning

IBM Roadrunner

32

System Name	Roadrunner			
Site	DOE/NNSA/LANL			
System Family	IBM Cluster			
System Model	BladeCenter QS22 Cluster			
Computer	BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband			
Vendor	IBM			
Application area	Not Specified			
Installation	Year 2008			
Operating System	Linux			
Interconnect	Infiniband			
Processor	PowerXCell 8i 3200 MHz (12.8 GFlops)			
Cores	Rmax(GFlops)	Rpeak(GFlops)	Nmax	Nhalf
129600	1105000	1456704	2329599	0

1GFlops = 10^9 = 1 000 000 000 de calculs en virgule flottante par seconde

De la puissance des machines ...

33



La PS3 est 800 fois plus puissante que le premier supercalculateur vectoriel Cray-1

Comparaison Cray 1 / PS3
sortie 1976 / 2005
80 MHz / 3,2 GHz

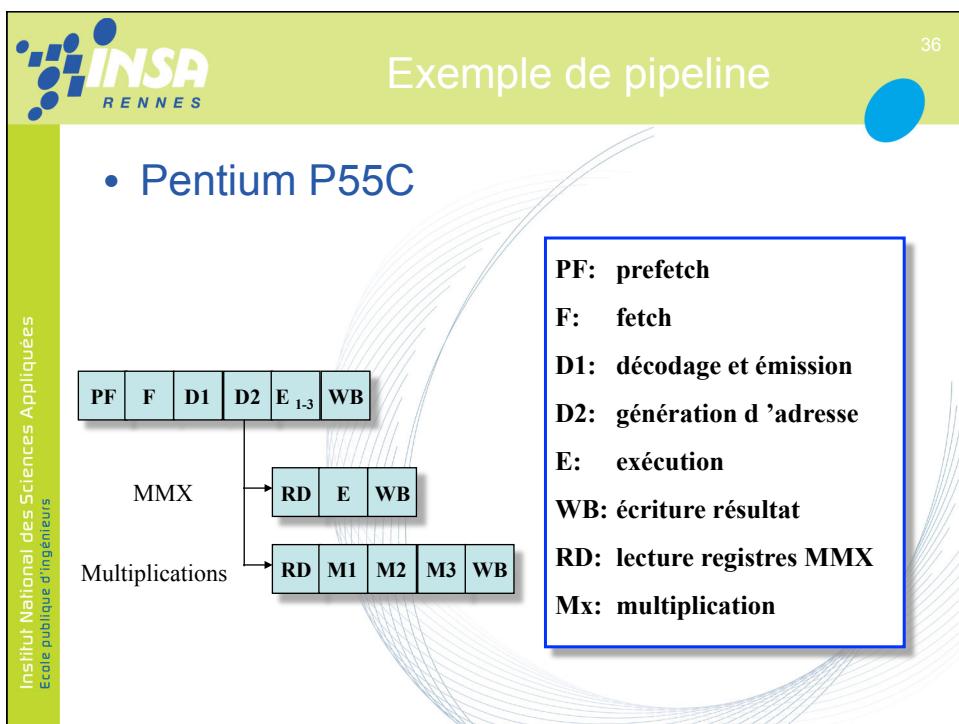
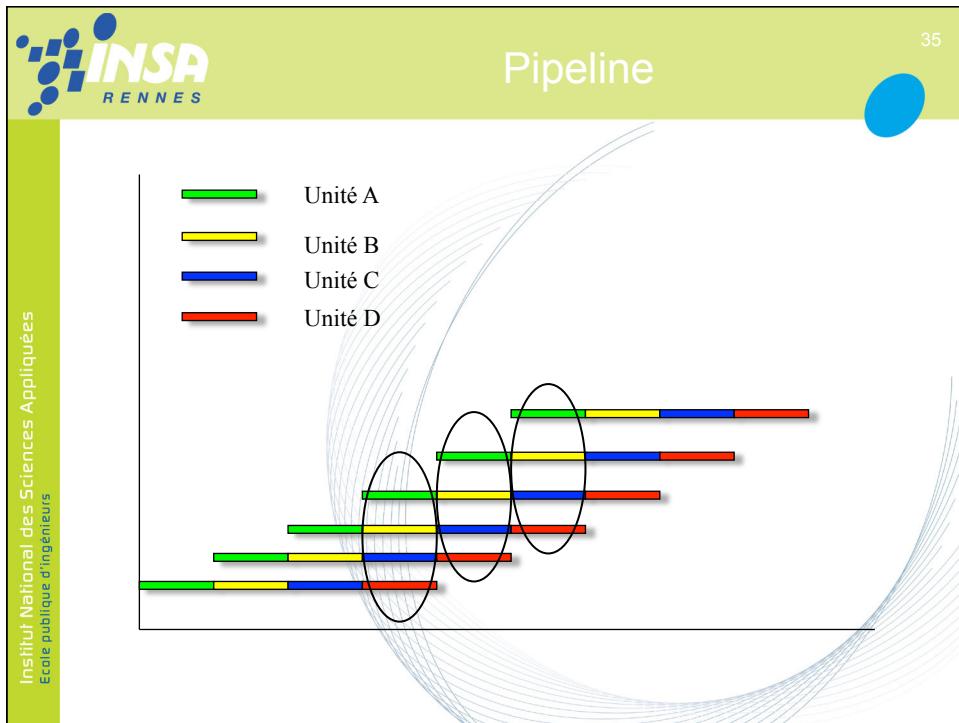
250 MFlops/200 Gflops théorique
coût \$5M/\$8M / \$500/\$600 au lancement
consommation 60kW / 200W



Parallélisme d'exécution

34

- Vu en cours système
 - pour les DMA, accès disques, ...
- Quelques types de machines
 - pipelines, calculateurs vectoriels (SIMD)
 - multiprocesseurs (MIMD)
 - à mémoire partagée
 - à mémoire distribuée
 - réseaux de machines (NOW / COW)

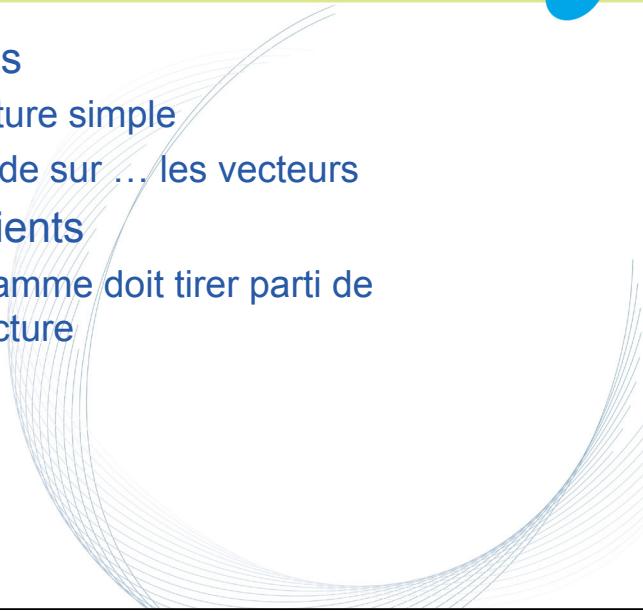


- Echanges mémoire <-> Proc
 - Vecteurs
- Calculs
 - Vecteurs
- Bien adaptés à la simulation numérique de déformations (crash tests)

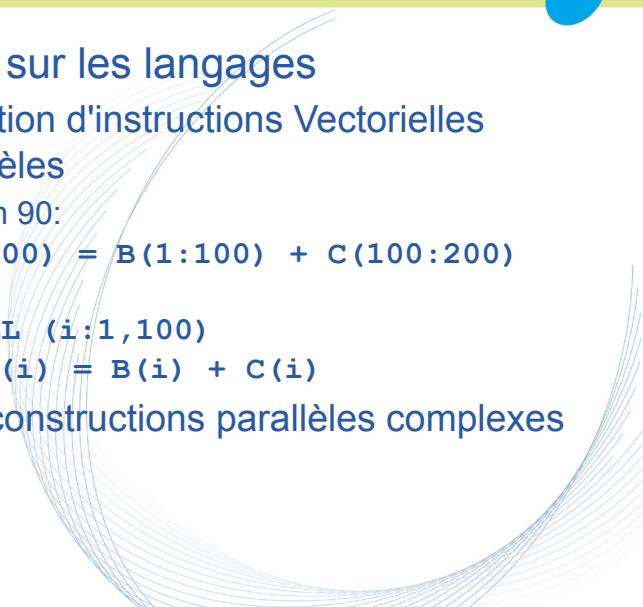
- Dernière série de calculateurs vectoriels de Cray research inc :
 - série T90
 - Sorti en 1995
 - 450 MHz 1,8 Gflops/proc
 - Jqa 32 procs
 - 1 module mémoire =256 Mo en 16 bancs mémoire statique
 - T94 8 Gflops
 - jqa 1 Go de mémoire, accès 100 Go/s
 - 2 unités vectorielles par processeur
 - registres vectoriels
 - Autre exemple: [NEC SX5](#)



- Avantages
 - architecture simple
 - très rapide sur ... les vecteurs
- Inconvénients
 - le programme doit tirer parti de l'architecture

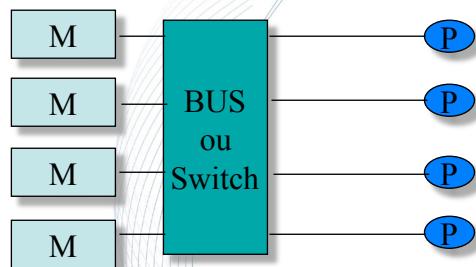


- Influence sur les langages
 - introduction d'instructions Vectorielles et parallèles
 - Fortran 90:
- A(1:100) = B(1:100) + C(100:200)**
- FORALL (i:1,100)**
A(i) = B(i) + C(i)
- pas de constructions parallèles complexes



Calculateurs à mémoire partagée

41



Calculateurs à mémoire partagée

42

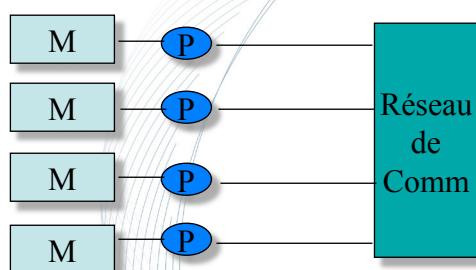
- Exemple SGI Challenge XL
 - 1,2 Go/s
 - 60 Mo à 16 Go
 - 2 à 32 processeurs superscalaires R10000
 - 5 instructions par cycle
 - 250 MHz (600 MHz)



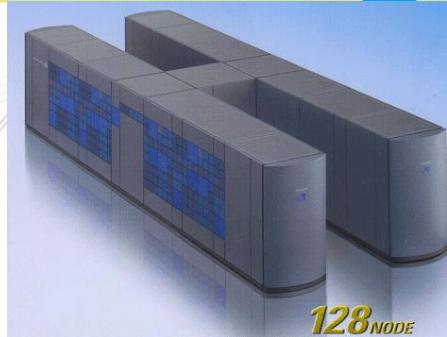
- Avantages
 - conception assez simple
 - utilisation facile
- Inconvénients
 - bande passante mémoire, conflits
 - coût de l'interconnexion
 - nombre de processeurs limité

- 2 sous groupes
 - SMP (Symetric multiprocessors)
 - PC Bi/Quadri processeurs
 - CC-NUMA
 - Cache Coherent Non Uniform Memory Access
 - Partage mémoire en apparence seulement
 - Exemple SGI origin 2000

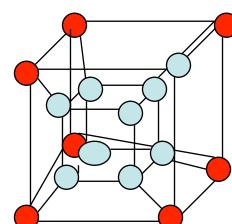
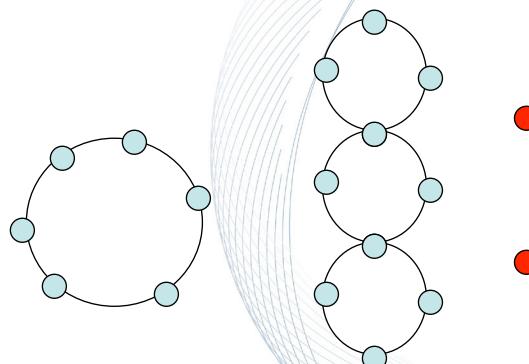
- Influence sur les langages
 - Notion de processus
 - espace d'adressage protégé
 - Notion de tâches légères
 - espace d'adressage partagé
 - Variables partagées
 - Synchronisations pour empêcher les conflits
 - Base : exclusion mutuelle



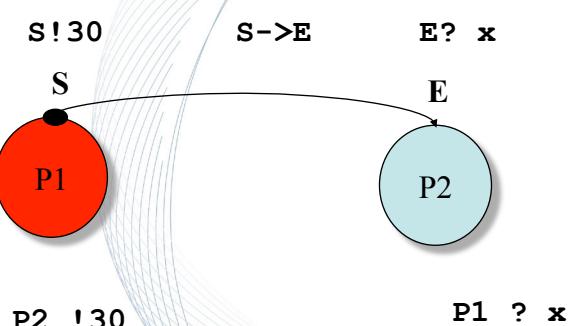
- Exemples:
 - iPSC1 & 2
 - intel Paragon
 - Hitachi SR 8000
 - jqa 128 procs
 - 8 Gflops/noeud
- espèce en voie de disparition (?)
- Renaissance avec les « grappes de PC »



- Topologie du réseau d'interconnexion



- Influence sur le langage
 - notion de processus
 - pas de partage de variables
 - notion de "message"
 - échange de valeurs
 - coopération explicite



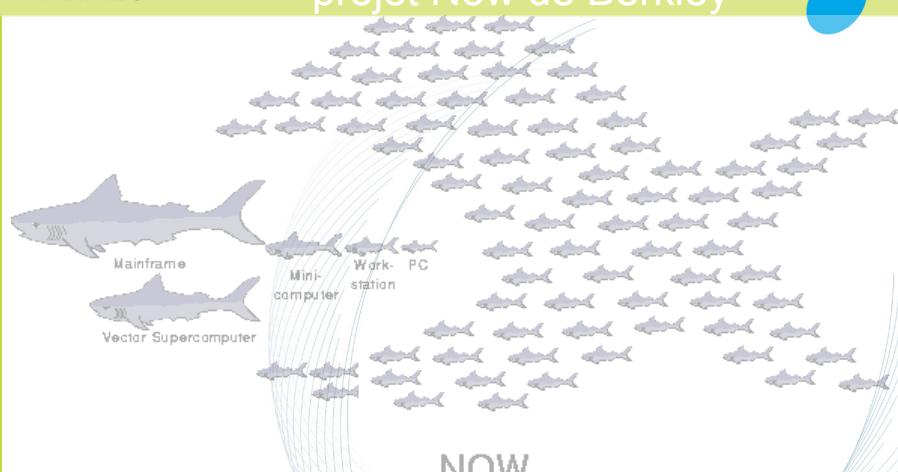
Calculateur « exotique » (machine SIMD)

- Connection machine
 - 16384 processeurs, 4 Ko par processeur
 - 1 seul séquenceur:
 - A chaque cycle un processeur fait soit
 - la même opération que les autres
 - rien
 - adapté au calcul vectoriel et matriciel
 - parallélisme de données
 - A influencé le design des GPU actuels
 - Video supercomputercooled

- Influence sur les langages
 - langages à parallélisme de données :
 - opérations globales sur les données
 - alpha opérations, beta réductions
 - déplacement de données (SHIFT, ROTATE, ...)

- Même architecture qu'un calculateur à mémoire distribuée
- Interconnexion de stations de travail
 - réseaux rapides
 - 100Mbit/s, 1 Gbit/s
- Facilement extensibles

Les stations peuvent aussi être des machines parallèles !



On April 30th, 1997, the NOW team achieved over 10 GFLOPS on the [LINPACK](#) benchmark, propelling the NOW into the top 200 fastest supercomputers in the world!

- Avantages
 - relativement faible coût
 - extensibles
 - matériel standard
- Inconvénients
 - puissance "relativement" modeste
 - réseaux d'interconnexion limité par leur interface sur la station

- Mémoire partagée
 - tâches légères
 - variable partagées
 - synchronisations
- Mémoire distribuée
 - processus
 - communication à distance

- Et les applications dans tout ça ?
 - Utiliser un modèle de programmation adapté
 - au modèle
 - transformer le programme pour l'exécuter (coûteux)
 - utiliser un système (exécutif) adapté (coûteux aussi)
 - à l'architecture
 - « contorsions » pour la programmation
 - portage sur une autre architecture difficile

- Introduction
- Parallélisme d'expression
 - Applications
- Parallélisme d'exécution
 - Architecture de machines
 - Influence sur l'expression du parallélisme
- Challenges

- Maîtriser
 - l'utilisation des mécanismes de synchronisation et de communication
 - les techniques de programmation parallèle
- Comprendre
 - la complexité des systèmes parallèles et distribués
 - les différents comportements parallèles

- Cette année
 - Synchronisation
 - Programmation parallèle
 - Programmation distribuée
- En 5e année
 - Algorithmique distribuée
 - Calcul numérique parallèle (?)