

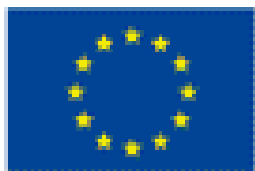
RICERCATORI...
GENTE CHE LASCIA IL SEGNO.

NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

Incontro dei Ricercatori dell'Area di Ricerca con gli studenti di Roma e dei Castelli

Antonio Perozziello
Fabrizio Poggi

ENEA UTICT HPC
ENEA UTICT HPC



Gli Enti di Ricerca nell'Area Romana

L' Area di Ricerca di Frascati nasce nel **1953** con la decisione di costruire un acceleratore di elettroni:
Il Sincrotrone.

La costruzione e le attività di ricerca erano gestite dal **CNEN (1952)** e **INFN (1951)**

Da allora l'area si è espansa con

ESRIN-ESA , CNR , INGV, ASI, Tor Vergata , Policlinico

....

NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

www.frascatiscienza.it



Gli Enti di Ricerca nell'Area Romana



NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

www.frascatiscienza.it



Frascati Scienza



L'area di Frascati rimane protagonista indiscussa, **da oltre 50 anni**, di importanti attività di ricerca di livello internazionale. Tanto da essere da molti definita la **città della scienza**.

L'**Associazione Frascati Scienza** nasce dall'esigenza di condividere e partecipare a questo importante patrimonio, **promuovendo l'educazione alla ricerca scientifica** e la sua comunicazione al largo pubblico, attraverso eventi di divulgazione e reti di comunicazione permanente fra ricercatori e cittadini.

La nostra **Area di Ricerca e Frascati Scienza** sono i protagonisti dell'evento **Notte Europea dei Ricercatori** sin dalla prima edizione del 2006

NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

www.frascatiscienza.it

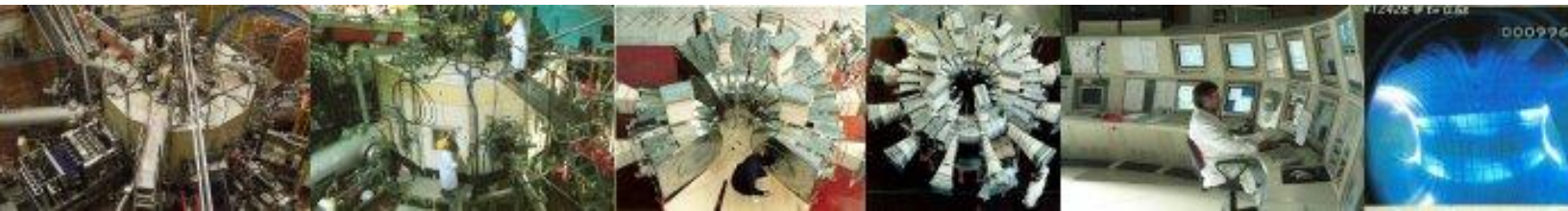


Venerdì 23 Settembre

La Notte Europea dei Ricercatori:

i laboratori vengono aperti al pubblico

- ✧ **CNR e INAF – Area di ricerca di Tor Vergata ARTOV**
 - ✧ Astronomia
 - ✧ Fisica cosmica
 - ✧ Astrofisica spaziale
 - ✧ Medicina e scienze della vita
- ✧ **ENEA (Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile)**
 - ✧ Ambiente e clima
 - ✧ Nucleare : Fusione e Fissione (Casaccia)
 - ✧ Fonti rinnovabili
 - ✧ Metodi fisici per i Beni Culturali



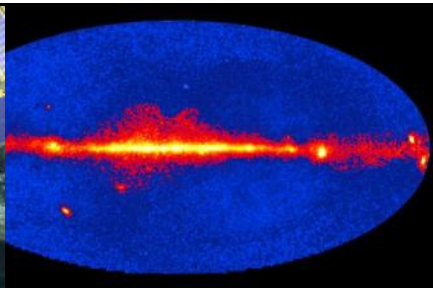
Venerdì 23 Settembre

La notte Europea dei Ricercatori:

i laboratori vengono aperti al pubblico

- ✧ **ESA ESRIN**
 - ✧ Terra e ambiente spaziale circostante
 - ✧ Sistema solare e Universo in generale
 - ✧ Servizi satellitari

- ✧ **ASI – ASDC (Agenzia Spaziale Italiana – Asi Science Data Center)**
 - ✧ Abitabilità nello spazio
 - ✧ Osservazione della terra ed esplorazione del sistema solare
 - ✧ Telecomunicazioni
 - ✧ Trasporto spaziale

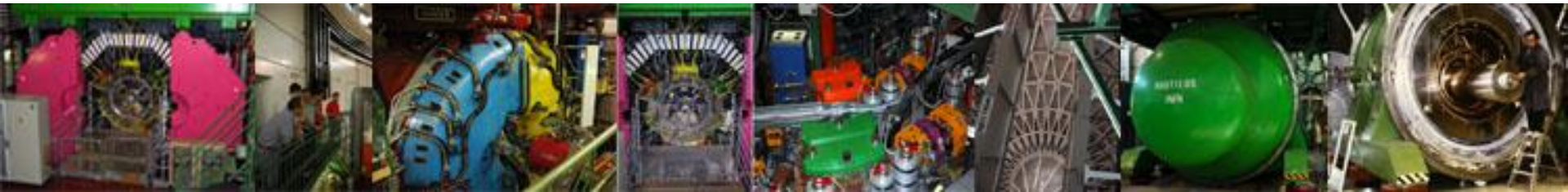


Venerdì 23 Settembre

La notte Europea dei Ricercatori:

i laboratori vengono aperti al pubblico

- ✧ **INFN – LNF (Istituto Nazionale di Fisica Nucleare – Laboratori Nazionali di Frascati)**
 - ✧ Rivelatori di Particelle
 - ✧ Acceleratori di particelle ad alta energia
 - ✧ Simmetria Materia - Antimateria
 - ✧ Onde Gravitazionali
- ✧ **INGV (Istituto Nazionale di Geofisica e Vulcanologia)**
 - ✧ Sismologia applicata all'ingegneria
 - ✧ Geochimica
 - ✧ Sismologia e tettonofisica
 - ✧ Geomagnetismo e Geofisica Ambientale
- ✧ **INAF OAR (Osservatorio Astronomico di Roma)**
 - ✧ Astronomia



Seminari a Frascati dal 19 al 22 Settembre

*Sala degli Specchi – Comune di Frascati dalle **18.30 alle 19.30***

Lunedì 19 : Lo Stato di cultura e il futuro del libro con l'e-book

Prof. Fulco Lanchester, Ordinario di Diritto Costituzionale all'Università di Roma Sapienza

Martedì 20 : Copernico e Galileo, oltre l'eliocentrismo

Prof. Paolo Simoncelli, Ordinario di Storia Moderna all'Università di Roma Sapienza

Mercoledì 21 : Tra Scienza e Filosofia

Prof.ssa Teresa Serra, Ordinario di Filosofia del Diritto all'Università di Roma Sapienza

Giovedì 22 : Teoria della complessità e scienze della vita:
oltre il riduzionismo e il determinismo

Dott. Mirko Di Bernardo, ricercatore presso l'Università di Roma Tor Vergata

NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

www.frascatiscienza.it



Venerdì 23 Settembre

La Notte Europea dei Ricercatori a Frascati

17.00~24.00 : Hands on experiments (*Passeggiata e Scuderie Aldobrandini*)

Giovani Laureati effettueranno piccoli esperimenti ed esibizioni dal vivo

17.00~24.00 : Star Party (*Passeggiata*)

Osservazioni Astronomiche lungo la passeggiata Belvedere di Frascati

21.30 ~23.00 :

La Chimica in Casa Spettacolo condotto da Riccardo Rossi (*Scuderie Aldobrandini*)

Spazio bimbi Esperimenti di biologia e laboratori per i più piccoli (*Passeggiata*)

Mostra fotografica “Natura, Energia e Sviluppo” (*Scuderie Aldobrandini*)

Vignette Stellari esposizione del vignettista frascatano Andy Ventura (*Scuderie Aldobrandini*)

Energy Songs Coro Giovanile Diapason (*Frascati centro*)

20.45~22.45 : Rappresentazione teatrale “Il kyoto fisso” (*Centro ENEA di Frascati*)

20.30~24.00 : La Notte europea dei ricercatori sotto il cielo di Roma (*Univ. Roma Tre*)

Conversazioni scientifiche, osservazioni astronomiche e attività hands on .

NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

www.frascatiscienza.it



Seminari nelle Scuole



Frascati Scienza mette a disposizione una rete di ricercatori di altissima competenza e la professionalità a livello mondiale.

L'incontro con il ricercatore di quest'anno mette in contatto **30 ricercatori con 20 Istituti Scolastici** di Roma e Castelli

L'associazione **Frascati Scienza** permette ai suoi soci di organizzare anche durante l'anno corsi e seminari presso le scuole del territorio.

Nel 2010 sono stati organizzati:

- **La Fisica per i Presidenti del Futuro** (corso di 10 ore)
- Seminari su problemi energetici e disastri ambientali

NOTTE EUROPEA DEI RICERCATORI
VENERDÌ 23 SETTEMBRE 2011
A FRASCATI Ingresso libero

www.frascatiscienza.it





INTRODUZIONE AL CALCOLO PARALLELO

Antonio Perozziello UTICT HPC
Fabrizio Poggi UTICT HPC

Calcolo sequenziale:

Risolve un problema tramite un **algoritmo**, le cui istruzioni sono eseguite in sequenza.

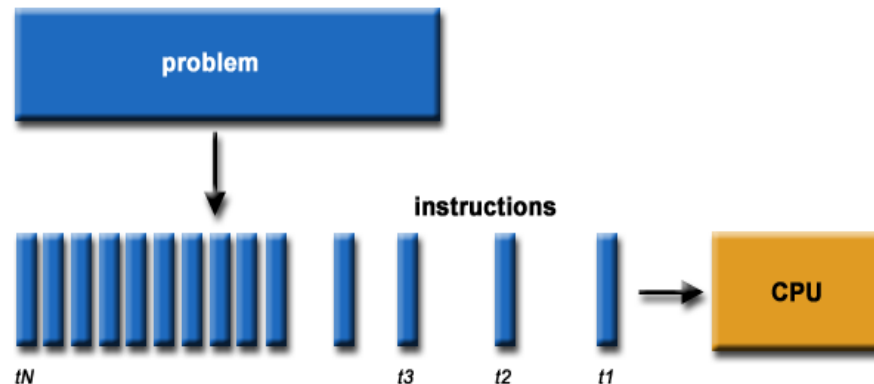
Un **algoritmo** è un procedimento per la risoluzione di una classe di problemi, costituito da un insieme finito di direttive non ambigue che specificano una sequenza finita di operazioni da eseguire su un insieme finito di dati.

Il **programma** è la descrizione di un algoritmo in una forma comprensibile al calcolatore. Si fa girare il programma su un solo computer con una sola CPU (Central Processing Unit)

Il programma è scritto con istruzioni successive.

Le istruzioni vanno eseguite una dopo l'altra.

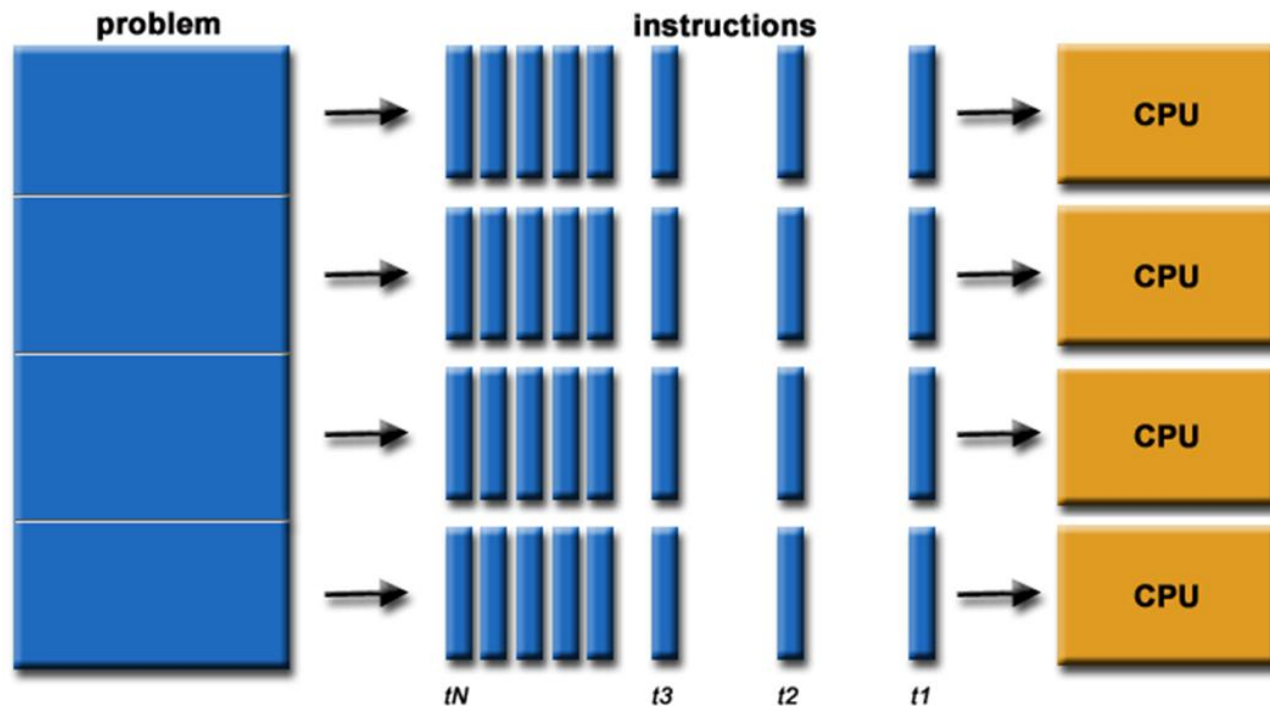
Solo una istruzione può essere eseguita nella stessa unità di tempo.



Calcolo parallelo (parallel computing)

Uso simultaneo di più risorse di calcolo utilizzate per risolvere un problema.

- 1) I programmi vengono suddivisi in moduli che sono eseguiti concorrentemente e/o sequenzialmente su più processori.
- 2) I dati possono essere ripartiti tra i processori e/o essere condivisi in una memoria comune.
- 3) I moduli in esecuzione sui diversi processori comunicano mediante messaggi per scambio di dati e sincronizzazione dell'attività.



Espressione aritmetica $a_1+a_2+a_3+a_4+a_5+a_6+a_7+a_8$

Esecuzione sequenziale

- 1) a_1+a_2
- 2) $(a_1+a_2)+a_3$
- 3) $(a_1+a_2+a_3)+a_4$
- 4) $(a_1+a_2+a_3+a_4)+a_5$
- 5) $(a_1+a_2+a_3+a_4+a_5)+a_6$
- 6) $(a_1+a_2+a_3+a_4+a_5+a_6)+a_7$
- 7) $(a_1+a_2+a_3+a_4+a_5+a_6+a_7)+a_8$

$p = 1$ (nr. Proc.)

$T = 7$ (nr. Unità di tempo)

Espressione aritmetica

$$a_1+a_2+a_3+a_4+a_5+a_6+a_7+a_8$$

Esecuzione parallela

- 1) a_1+a_2 a_3+a_4 a_5+a_6 a_7+a_8
- 2) $(a_1+a_2)+(a_3+a_4)$ $(a_5+a_6)+(a_7+a_8)$
- 3) $(a_1+a_2+a_3+a_4)+(a_5+a_6+a_7+a_8)$

$$p = 4 \quad (\text{nr. Proc.})$$

$$T = 3 \quad (\text{nr. Unità di tempo})$$

QUINDI

- per essere reso parallelo, il problema computazionale deve essere tale che:
 - Può essere spezzato in parti di lavoro che possono essere risolte contemporaneamente;
 - Possono essere eseguite istruzioni multiple del programma in ogni frazione di tempo;
 - Può essere risolto in minor tempo con risorse di calcolo multiple piuttosto che con una singola risorsa di calcolo.

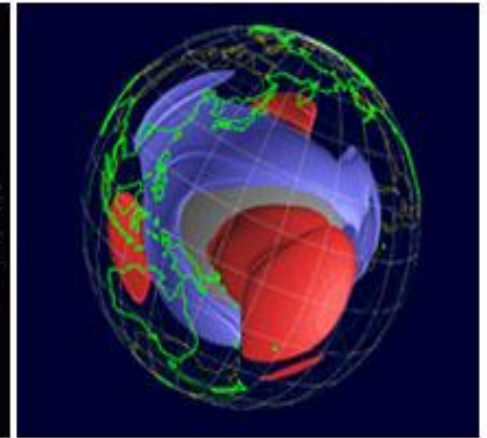
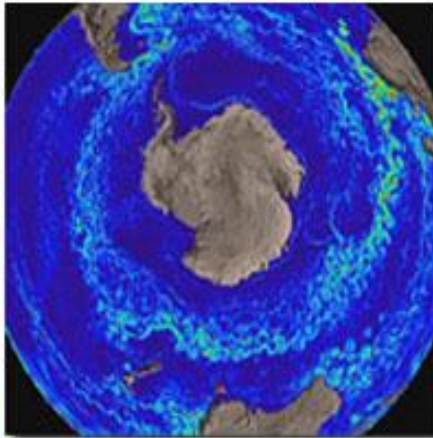
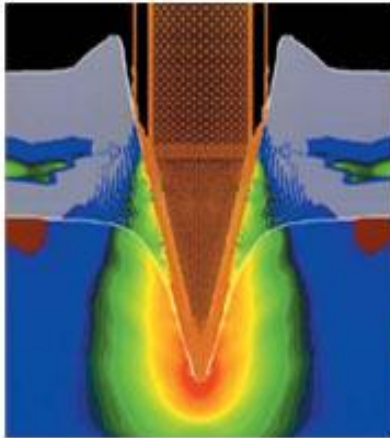
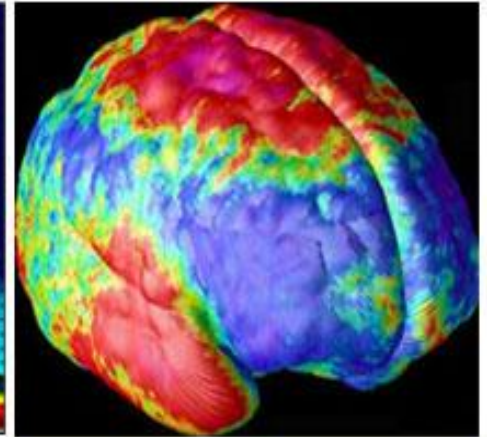
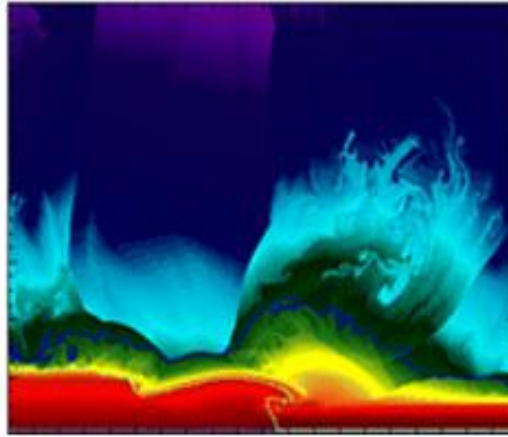
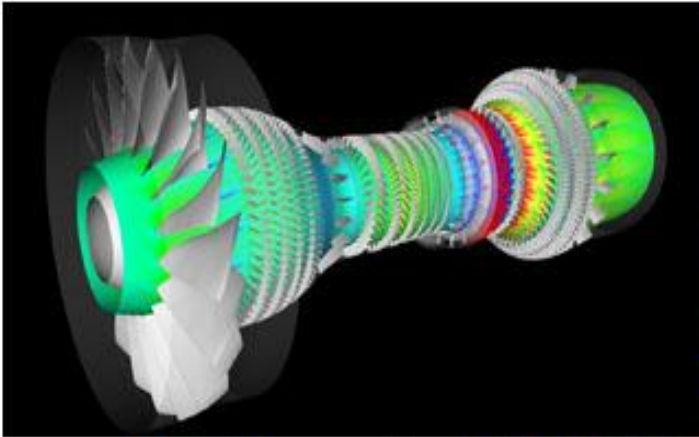
Le risorse di calcolo per il parallelo possono essere:

- **Un singolo computer con più processori**
- **Un numero arbitrario di computer connessi da una rete**
- **Una combinazione delle due cose**

- Calcolo parallelo
una evoluzione del calcolo seriale che tenta di emulare quanto avviene nel mondo: molti eventi complessi, che accadono allo stesso tempo, all'interno di una sequenza.
- Esempi:
 - Formazione delle galassie, movimenti planetari, linee di assemblaggio delle autovetture, traffico cittadino, modelli atmosferici, costruzione di una navicella spaziale, ecc.....

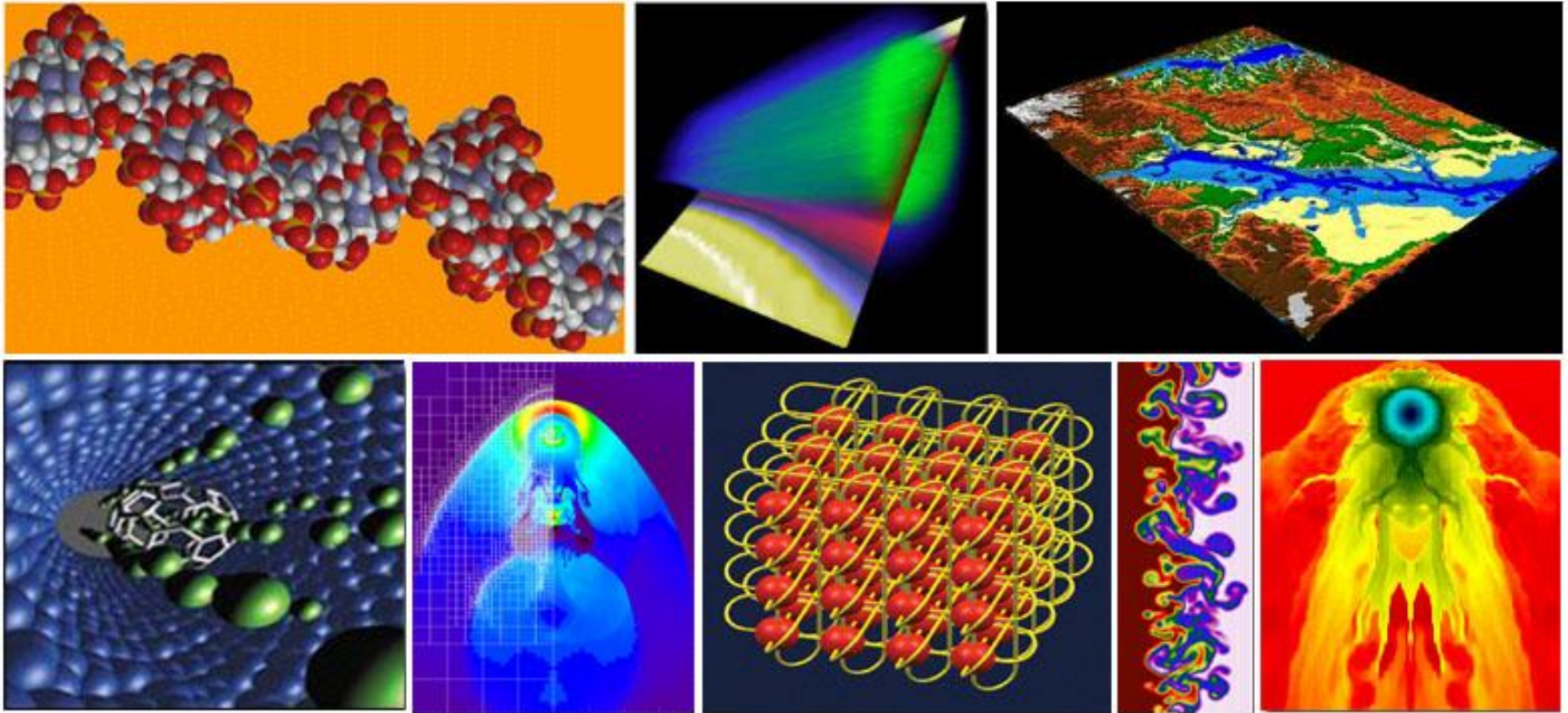


- Il calcolo parallelo viene dunque utilizzato per modelli scientifici difficili e problemi di ingegneria che si incontrano nel mondo reale. Per esempio:
 - Atmosfera, terra, ambiente; fisica applicata, nucleare, particelle, alta pressione, fusione, fotonica; bioscienze, biotecnologie, genetica; Chimica, scienze molecolari; geologia, scienze molecolari; Ingegneria meccanica; Microelettronica; Informatica, matematica



Oggi anche le applicazioni commerciali forniscono un elemento trainante nello sviluppo di computer più veloci:

Banche dati, ricerche petrolifere, dispositivi di ricerca su Web, Immagini e diagnosi mediche, progettazioni farmaceutiche, grafica avanzata e realtà virtuale, particolarmente nell'industria del divertimento, tecnologie video e multi-media, dispositivi di lavoro a distanza.



- Il ***calcolo parallelo*** è l'esecuzione simultanea di un codice (diviso e specificamente adattato) su più microprocessori, o più **core** dello stesso processore, allo scopo di aumentare le prestazioni del sistema.
- Nonostante un sistema con un numero n di processori sia meno efficiente di un sistema con un singolo processore di velocità n volte superiore, esso è più economico.

- **Non si deve pensare che si possa ottenere un calcolo parallelo efficiente semplicemente mettendo più processori uno a fianco dell'altro e connettendoli ad una velocità sufficiente.**
- **La maggior parte degli algoritmi deve essere riscritta per poter sfruttare l'hardware parallelo. Un programma che viene svolto correttamente da una singola CPU potrebbe presentare problemi se svolto in parallelo: più copie dello stesso potrebbero interferire tra loro, ad esempio accedendo allo stesso indirizzo di memoria nella stessa frazione di tempo. Da qui la necessità di una attenta programmazione per sfruttare questo tipo di sistemi.**

Perché parallelo ???

Una singola risorsa di calcolo può fare una sola cosa per volta.

Il calcolo parallelo fa risparmiare tempo e denaro, poiché risorse di calcolo multiple, dedicate alla soluzione di un problema, accorciano il tempo di soluzione: infatti si riescono a fare molte cose simultaneamente. E i componenti di un sistema di calcolo parallelo possono essere processori di prezzo non eccessivamente alto.

- **Si possono risolvere problemi più vasti e difficili**, difficilmente risolvibili da un singolo computer, specialmente per la limitatezza della memoria indirizzabile di un singolo processore.
- Esempi:
 - Le ricerche sulle banche dati del Web che processano milioni di transazioni al secondo.
 - Oppure i problemi di **Grand Challenge** che richiedono PetaFlops e PetaBytes di risorse di calcolo.

- Un **grand challenge** (**grande sfida**) è un problema fondamentale nelle scienze e nell'ingegneria, con vaste applicazioni, le cui soluzioni potranno essere rese possibili con l'utilizzo di risorse di calcolo ad alte prestazioni disponibili nel futuro immediato.
- Esempi di grand challenges sono:
- **Fluidodinamica computazionale per**
 - La progettazione di aeromobili ipersonici, carrozzerie di automobili efficienti, e sottomarini estremamente manovrabili,
 - Previsioni del tempo con effetti a brevi e lunghi effetti,
 - Ritrovamento di stazioni petrolifere ad alto rendimento, e molte altre applicazioni;
- **Calcoli di strutture elettroniche per la progettazione di nuovi materiali come**
 - Catalizzatori chimici,
 - Agenti immunologici,
 - superconduttori;
- **Dinamiche del plasma per le tecnologie di energia a fusione e per la sicurezza e la tecnologia militare efficiente;**
- **Calcoli per capire la natura fondamentale della materia, includendo la cromo-dinamica quantistica e la teoria della materia concentrata;**
- **Calcoli simbolici come**
 - Riconoscimento di parole,
 - Visione computerizzata,
 - Comprensione di linguaggio naturale,
 - Modo di ragionare automatico,
 - Metodi per progettazione, produzione industriale e simulazione di sistemi complessi.

Con il calcolo parallelo **possiamo**, inoltre, **utilizzare risorse remote**, su una rete estesa (wide area), quando le nostre risorse locali sono scarse. La Università di Stanford, per esempio, utilizza circa 340.000 computer per una potenza di calcolo di 4.2 PetaFlops.

Inoltre, **ragioni fisiche e pratiche** pongono limiti significativi alla costruzione di computer sequenziali più veloci: la tecnologia dei processori sta portando ad un crescente numero di transistor da collocare su un chip. Tuttavia prima o poi si raggiungerà un limite estremo alla miniaturizzazione dei componenti.

La **velocità di un computer** è direttamente dipendente a quanto velocemente i nostri dati possono transitare all'interno dell'hardware.

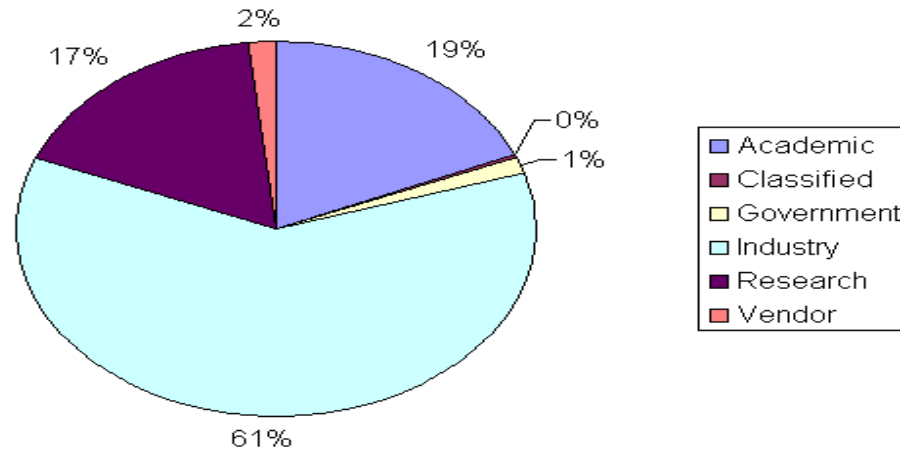
Limiti assoluti sono la velocità della luce (30 cm/nanosecondo) e il limite di trasmissione del filo in rame (9 cm/nanosecondo).

Diventa dunque particolarmente costoso fare un processore singolo più veloce.

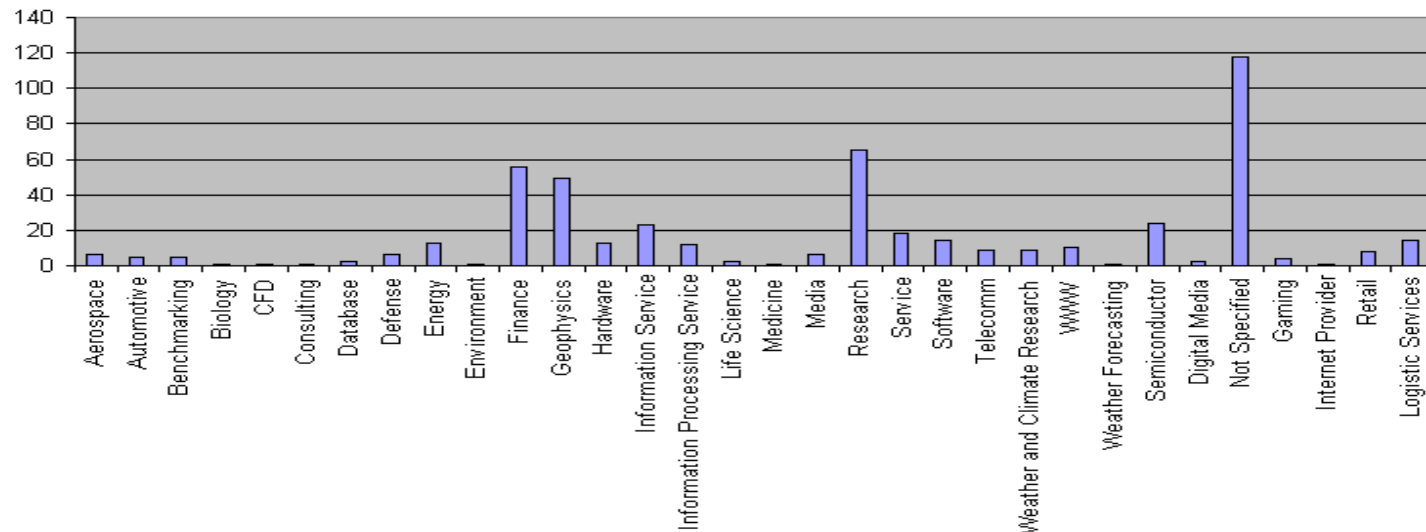
L'utilizzo di un gran numero di processori moderatamente più veloci, per raggiungere le stesse performance, è meno costoso.

Chi fa calcolo parallelo ?

Who's Doing Parallel Computing?

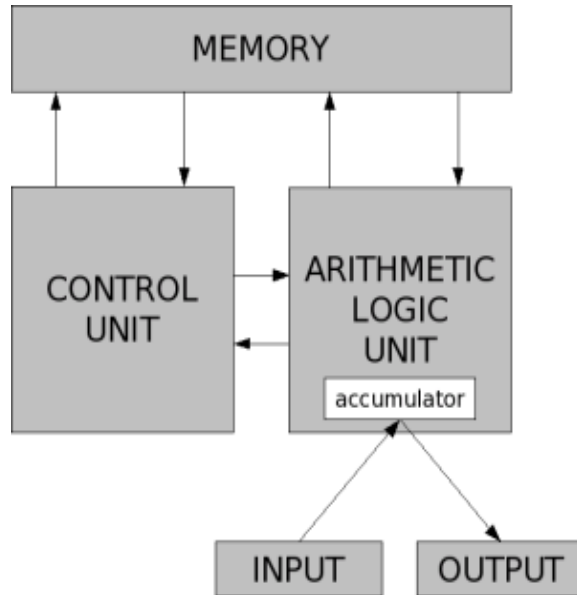


What Are They Using it For?



- **Il futuro:**
- Durante gli ultimi 20 anni, la tendenza verso reti sempre più veloci, sistemi distribuiti, architetture di computer a multi-processore (anche a livello di semplici desktop), mostrano chiaramente che il parallelismo è il futuro del calcolo scientifico e non solo.

- **Architettura di von Neumann**
- John von Neumann fu il primo che descrisse i requisiti generali per un computer (1945).
- Fin da allora, virtualmente tutti i computer hanno seguito questo schema di base.



Macchina di Von Neumann

E' composto da quattro componenti principali:

- Memoria
- Unità di controllo
- Unità Aritmetico-logica
- Input/Output

La memoria ad accesso Random di lettura e scrittura è usata per memorizzare le istruzioni di programma e i dati.

- Le istruzioni di programma sono informazioni codificate che dicono al computer di fare qualcosa
- I dati sono semplicemente informazioni usate dal programma.

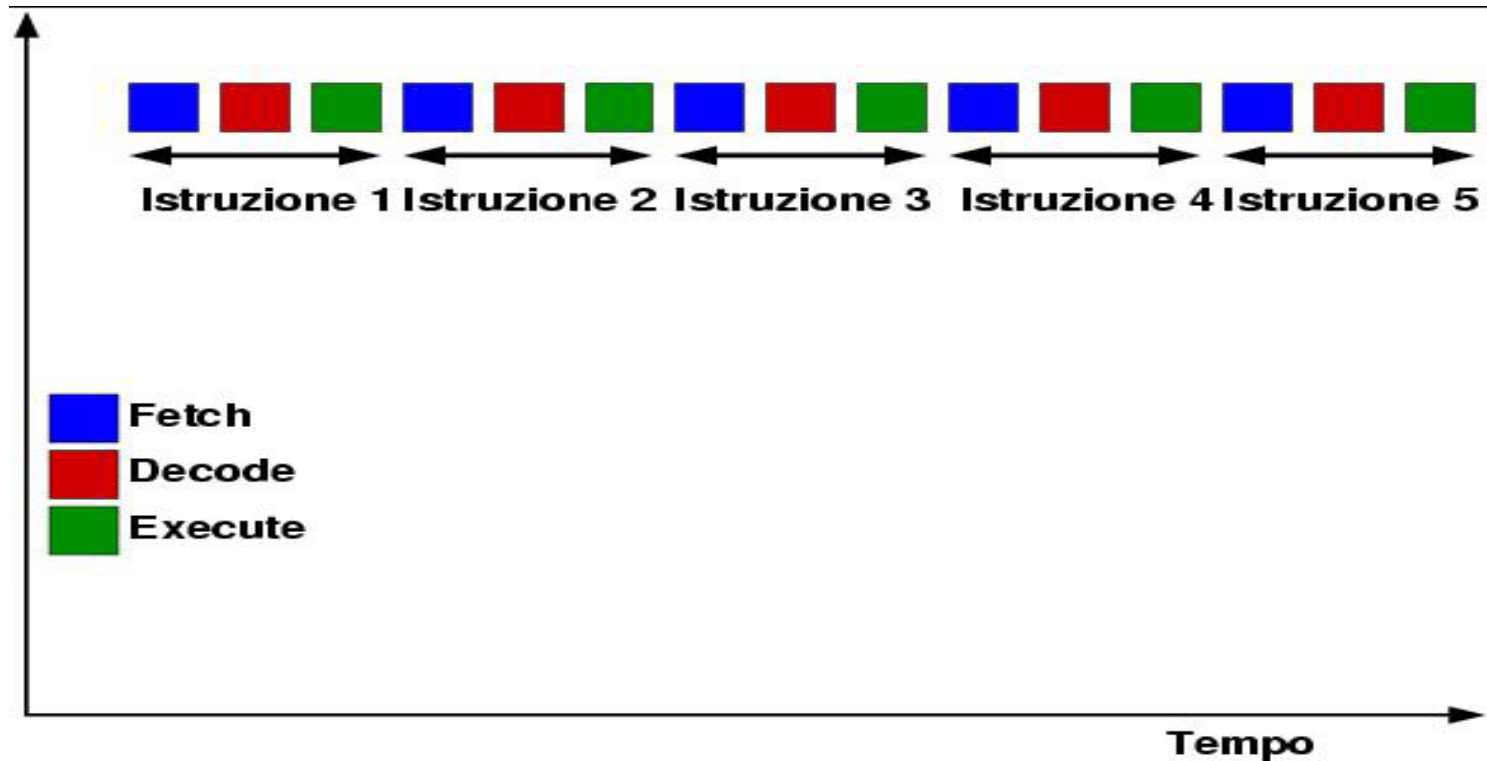
L'unità di controllo preleva le istruzioni e/o i dati dalla memoria, decodifica le istruzioni e poi **sequenzialmente** coordina le operazioni per completare il lavoro programmato.

L'Unità Aritmetica esegue le operazioni aritmetiche/logiche di base

L'Input/Output sono i dispositivi di interfaccia con l'operatore

- **I passi da fare per eseguire un'istruzione:**
 1. Prendere l'istruzione dalla memoria
 2. Decodificare l'istruzione
 3. Inizializzare i registri e prendere i dati dalla memoria
 4. Eseguire l'istruzione (per esempio: somma di numeri reali)
 5. Scrivere i risultati nella memoria.

CPU Sequenziale



Ogni quadrato → un ciclo di clock

Un'istruzione è composta da tre passi

Ci sono diversi modi per classificare i computer, soprattutto i computer paralleli. Una delle classificazioni più usata, in uso dal 1966, è chiamata Tassonomia di Flynn. La tassonomia di Flynn distingue le architetture dei computer, soprattutto quelli a multi-processore, a seconda di come esse possono essere classificate per le due dimensioni di Istruzioni e Dati. Ognuna di queste due dimensioni può avere solo uno dei seguenti possibili stati: singolo e multiplo. La tabella sotto definisce le quattro possibili classificazioni secondo Flynn:

SISD

Single Instruction, Single Data

SIMD

Single Instruction, Multiple Data

MISD

Multiple Instructions, Single Data

MIMD

Multiple Instructions, Multiple Data

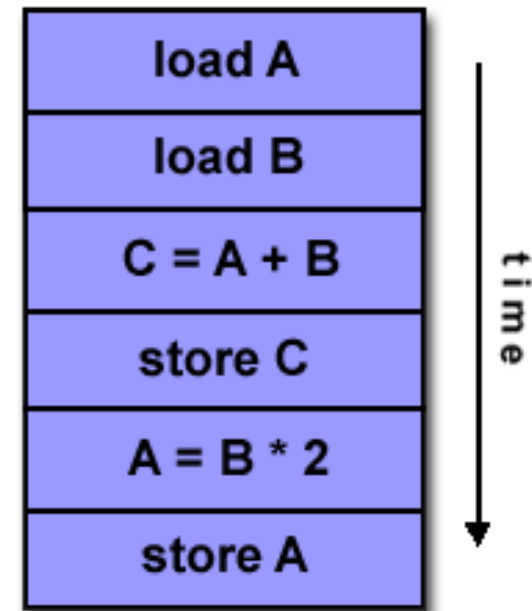
Single Instruction, Single Data (SISD):

Computer sequenziale. Istruzione singola: solo una istruzione macchina viene eseguita dalla CPU durante un ciclo di clock.

Dato singolo: solo un singolo dato viene usato come input durante un ciclo di clock.

E' il più antico e ancora oggi il tipo più comune di computer

Esempi: mainframes di generazione più vecchia, minicomputers, la maggior parte dei PC attuali.



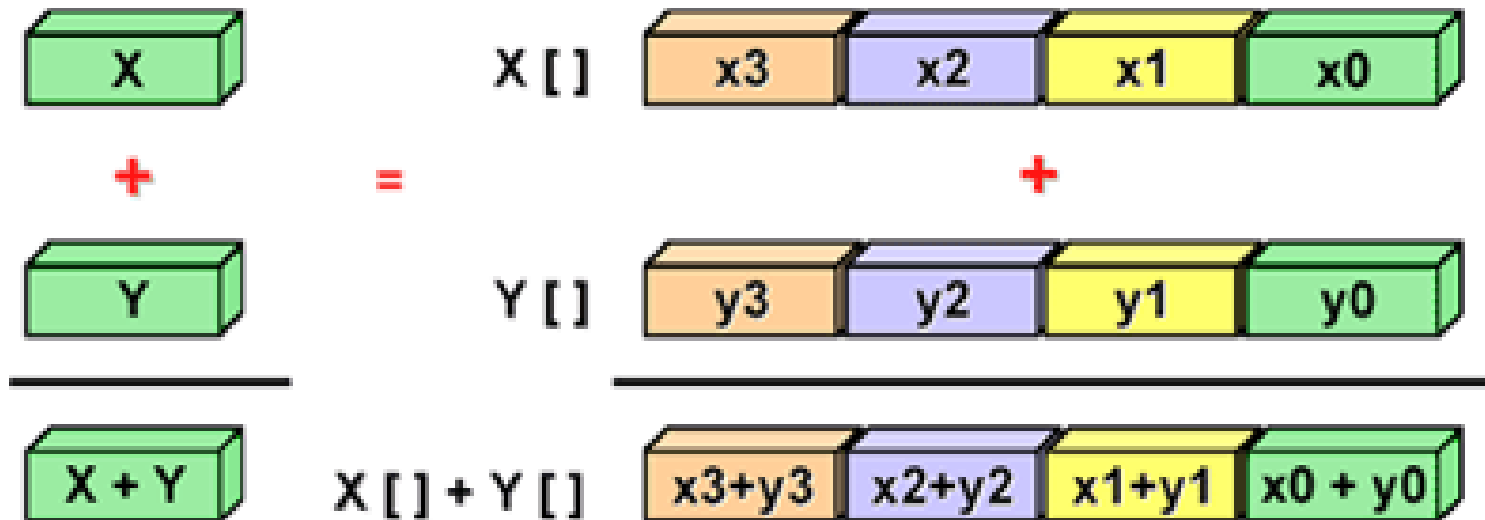
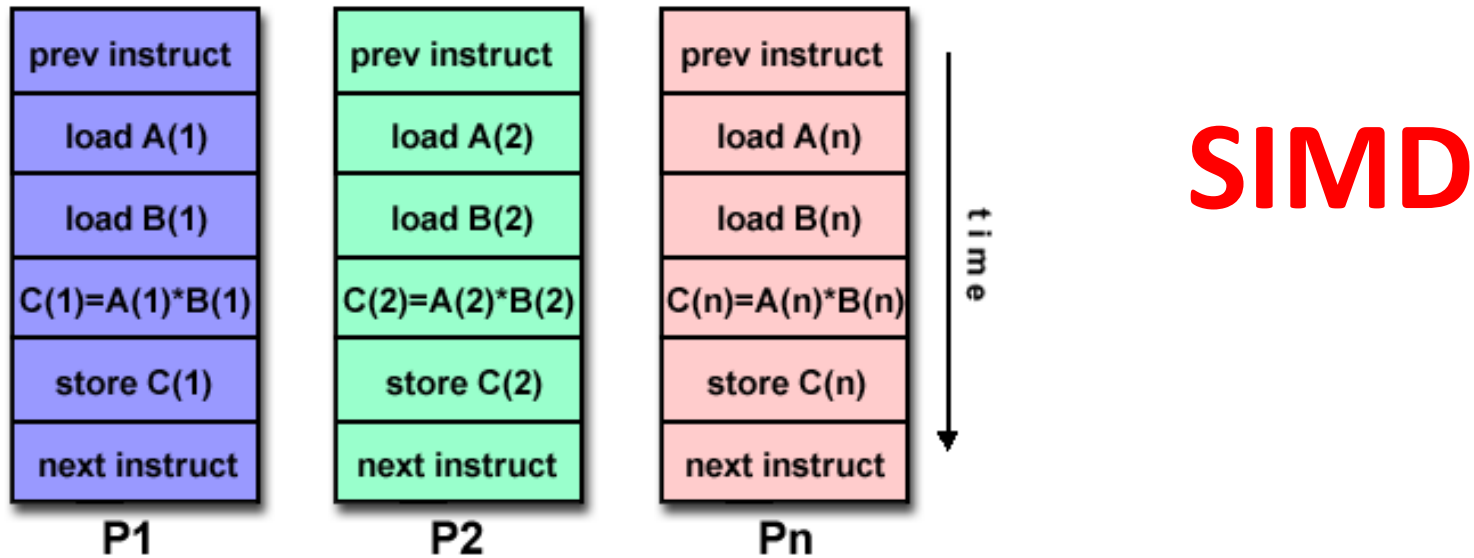


IBM 360



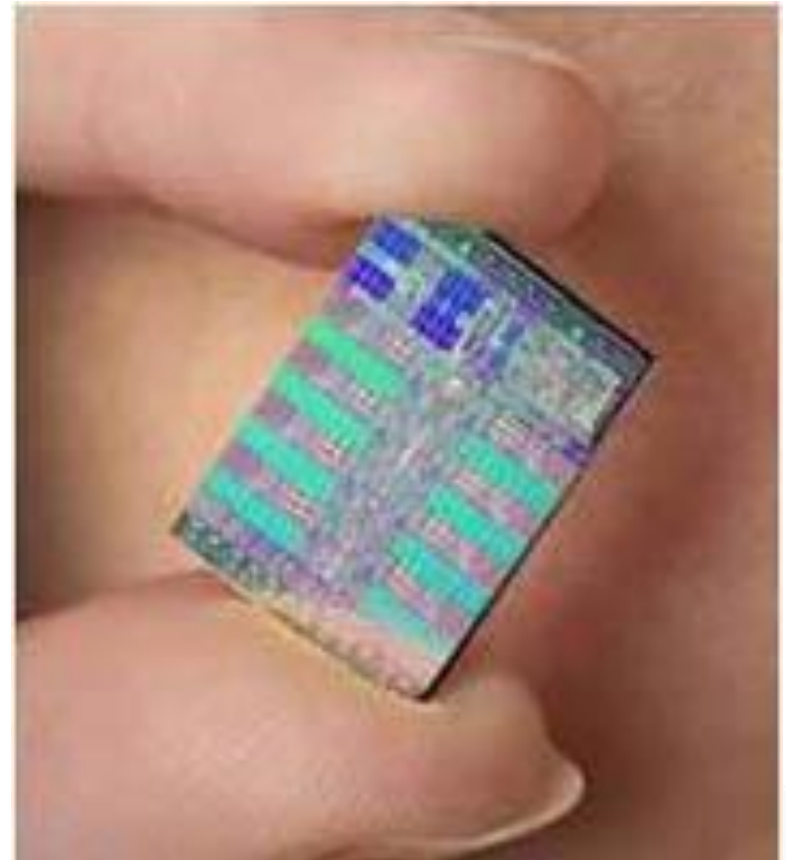
PDP 11

- **Single Instruction, Multiple Data (SIMD):**
- Un tipo di computer parallelo.
- **Istruzione singola:** tutte le unità di processo eseguono la stessa istruzione ad ogni ciclo di clock.
- **Dati multipli:** Ogni unità di processo può operare su un differente elemento di dati.
- Adatto a problemi specializzati, soprattutto per elaborazione di immagini.
- La maggior parte dei computer moderni, in particolare quelli con GPU (**Graphics Processor Units**), utilizzano unità di istruzione ed esecuzione SIMD.



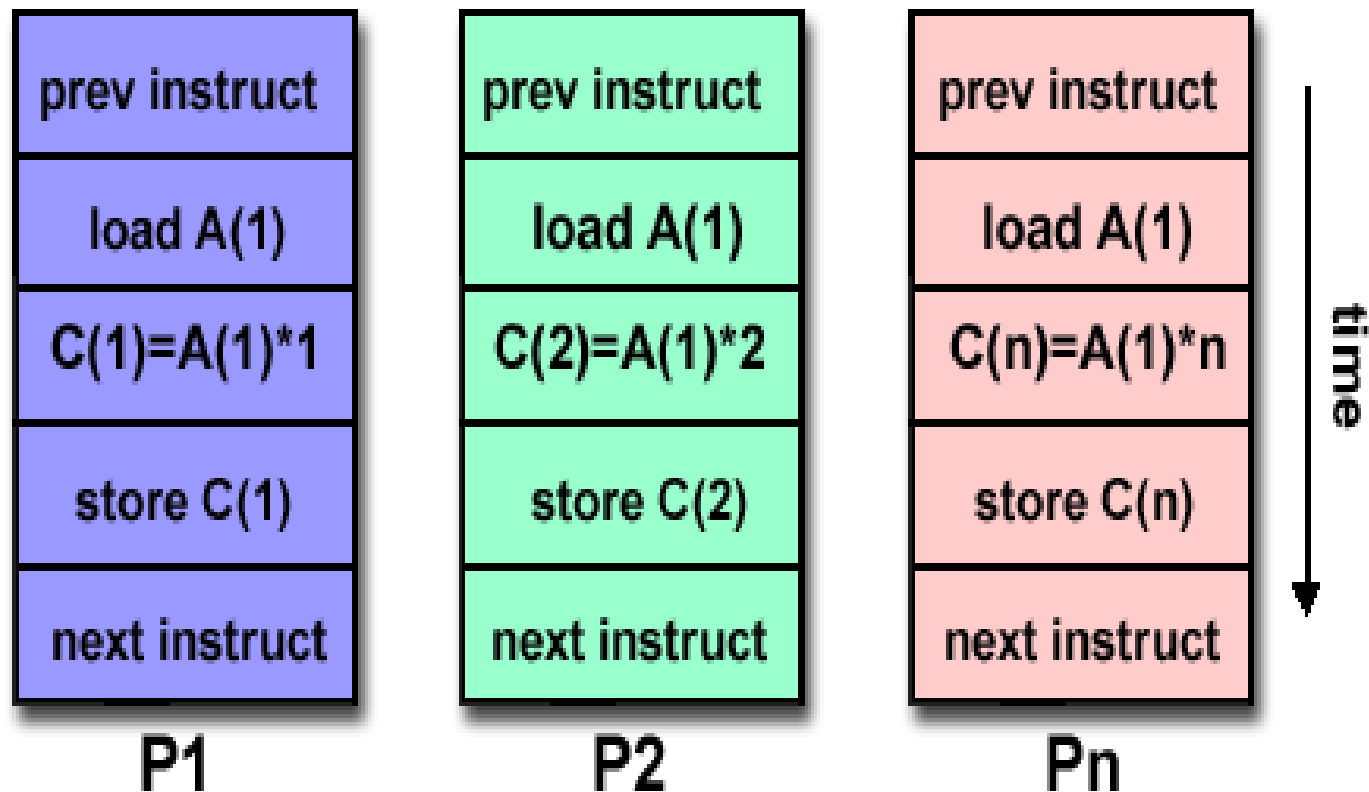


Cray X-MP



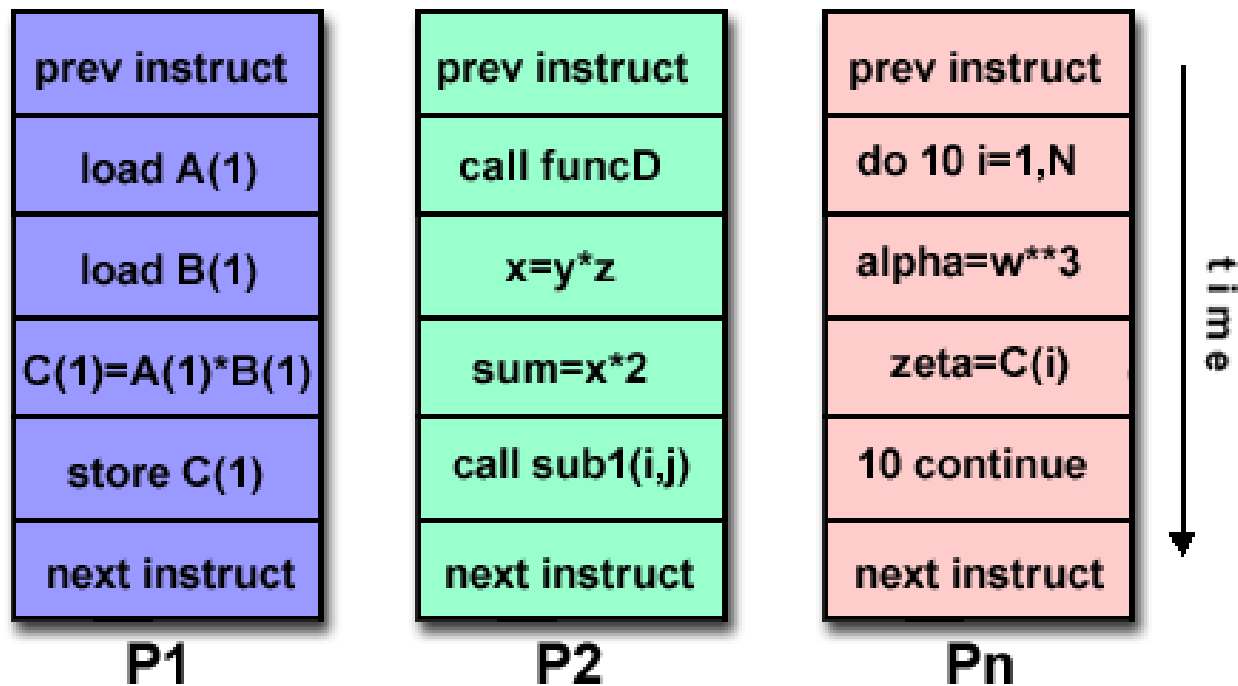
Cell Processor (GPU)

- **Multiple Instruction, Single Data (MISD):**
- Un singolo dato è immesso in unità multiple di calcolo.
- Ogni unità di calcolo opera sul dato attraverso flussi di istruzioni indipendenti.
- Ci sono pochi esempi di questa classe di computer paralleli. Da menzionare sono i computer che controllano il volo degli Shuttle spaziali.
- Un plausibile utilizzo potrebbe essere dato da:
algoritmi multipli di crittografia che tentano di decifrare un singolo messaggio codificato.



Un singolo dato, operazioni diverse

- **Multiple Instruction, Multiple Data (MIMD)**
- Il tipo più comune di computer parallelo. La maggior parte dei computer moderni ricade in questa categoria
- **Istruzioni multiple**: ogni processore può eseguire un flusso di istruzioni diverso
- **Dati multipli**: ogni processore può lavorare con un differente flusso di dati
- L'esecuzione può essere sincrona o asincrona.
- Esempi: la maggior parte dei supercomputers attuali, cluster e grid di computer paralleli su rete, computer SMP multi-processore, PC multi-core.





IBM POWER5



AMD Opteron

Terminologia

BIT (b)

Due soli stati (0,1), è l'unità minima di informazione

BYTE (B)

Insieme ordinato di 8 bit

$2^8 = 256$ combinazioni possibili (valori rappresentabili)



MSB

LSB

Terminologia

WORD

Numero byte (dipende dall'architettura del calcolatore)

1 byte	1 byte
--------	--------

Word da 16 bit

1 byte	1 byte	1 byte	1 byte
--------	--------	--------	--------

Word da 32 bit

1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
--------	--------	--------	--------	--------	--------	--------	--------

Word da 64 bit

- Terminologia:

Kilo	10^3	=	2^{10}
Mega	10^6	=	2^{20}
Giga	10^9	=	2^{30}
Tera	10^{12}	=	2^{40}
Peta	10^{15}	=	2^{50}
Exa	10^{18}	=	2^{60}

- Megasecondo** 1.000.000 sec = 11 giorni
- Gigasecondo** 1.000.000.000 sec = 31 anni

- Indirizzamento della memoria
- E' possibile indirizzare la memoria dall'indirizzo 0000.....00000 fino all'indirizzo 1111.....11111

il che significa:

per architetture a 32 bit $2^{32} = 4 \text{ GB}$

Per architetture a 64 bit $2^{64} = 16.000 \text{ PB}$

E' possibile classificare le macchine parallele rispetto alla gestione dell'indirizzamento della memoria:

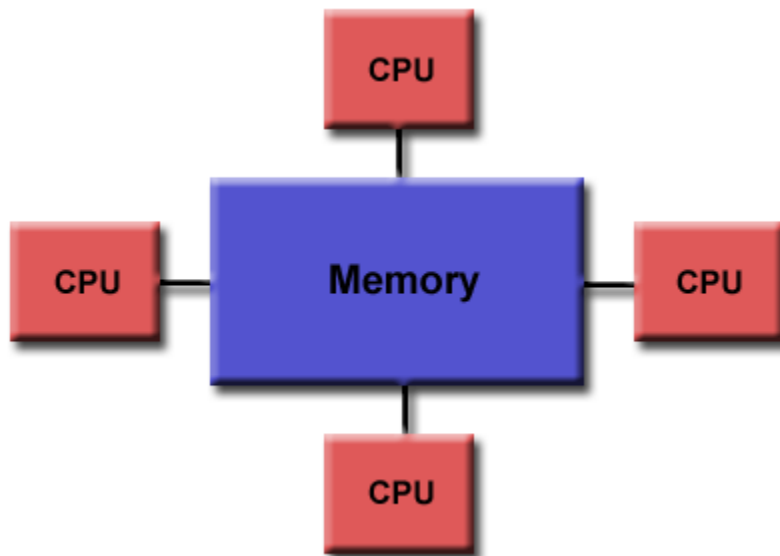
- 1) Shared Memory
- 2) Distributed Memory

- Architetture di computer paralleli rispetto alla memoria:
- 1) Shared Memory
 - Ogni processore può accedere tutta la memoria come uno spazio globale di indirizzi.
 - I processori operano indipendentemente condividendo le stesse risorse di memoria.
 - Un cambiamento in una locazione di memoria dovuto ad un processore è visibile a tutti gli altri.
 - Le macchine a memoria condivisa possono essere divise in due classi: **UMA** and **NUMA**.

MACCHINE A MEMORIA CONDIVISA

Uniform Memory Access (UMA):

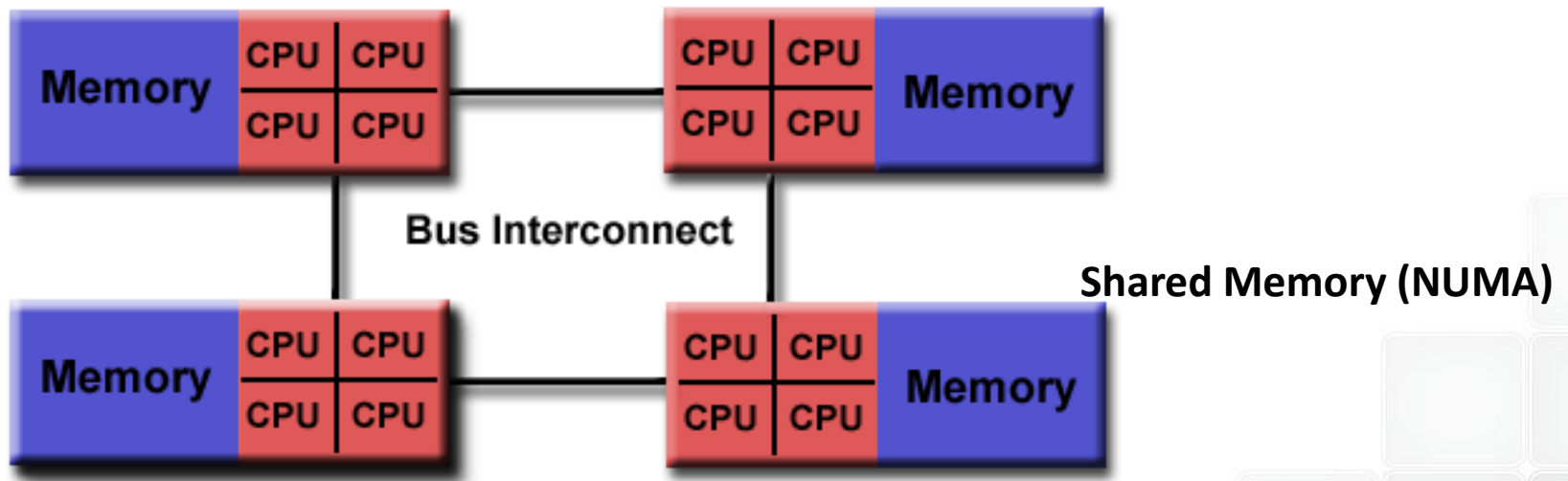
- Chiamati anche SMP (Symmetric Multiprocessor)
- Elaboratori costituiti da più processori identici
- I tempi di accesso a tutta la memoria sono uguali per ogni processore
- Se un processore aggiorna una locazione in memoria condivisa, tutti gli altri processori sanno dell'aggiornamento.



Shared Memory (UMA)

Non-Uniform Memory Access (NUMA):

- Sono realizzate attraverso il collegamento di due o più SMP
- Ogni SMP può accedere alla memoria degli altri SMP
- I processori non hanno lo stesso tempo di accesso a tutta la memoria



Shared memory

Vantaggi:

- Programmazione più agevole rispetto agli accessi alla memoria

Svantaggi:

- Scalabilità limitata
- E' compito del programmatore sincronizzare gli accessi alla memoria
- Aumentano i costi all'aumentare dei processori: diventa difficile e costoso progettare e produrre macchine a memoria condivisa con un numero troppo grande di processori (max 16).

Distributed Memory

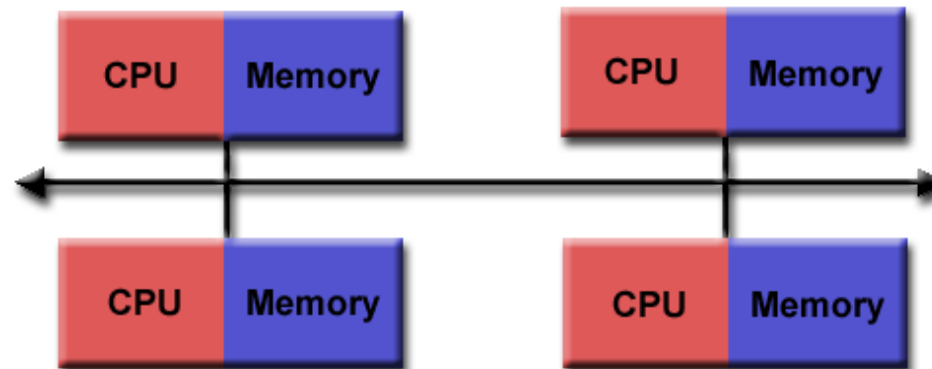
–Caratteristiche generali:

Ogni processore dispone di un'area di memoria locale

I processori possono scambiare dati solo attraverso il network di comunicazione

Non è previsto un concetto di spazio di indirizzi globale tra i processori: il programmatore dovrà definire in modo esplicito come e quando comunicare i dati e dovrà sincronizzare i task residenti su nodi diversi

E' importante l'infrastruttura di comunicazione utilizzata



Distributed memory

- **Vantaggi:**

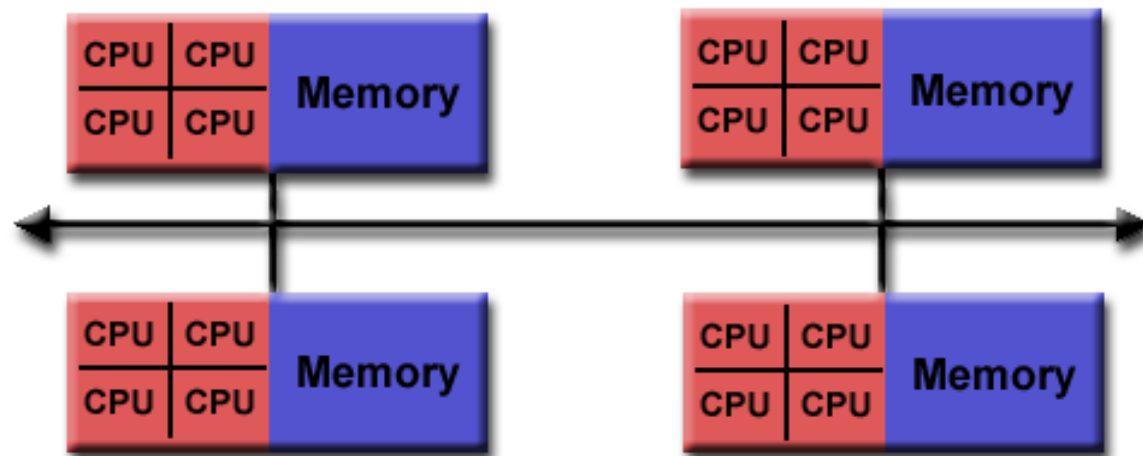
- La quantità di memoria è scalabile con il numero dei processori. Quando cresce il numero dei processori, cresce proporzionalmente la capacità della memoria.
- Ogni processore accede rapidamente alla sua propria memoria senza interferenza.
- Costo contenuto: sono più semplici da realizzare.

- **Svantaggi:**

- Il programmatore è responsabile di tutti i dettagli associati alla comunicazione di dati tra i vari processori
- Potrebbe essere complesso “mappare” codici esistenti su un modello di memoria distribuita
- E' cruciale la interconnessione

Hybrid Distributed-Shared Memory

- Il singolo nodo è un sistema shared memory
- I nodi sono interconnessi tramite una rete di comunicazione
- La comunicazione tra i processori avviene
 - tramite un'area di memoria condivisa (se sullo stesso nodo)
 - tramite la rete (se su nodi diversi)
- Questo tipo di architettura di memoria continuerà a essere il più usato negli anni futuri.



Interconnessione di rete

Rispetto alle macchine seriali, per il calcolo parallelo è estremamente importante (e costosa) la interconnessione

Nome	Latenza (microsecondi)	Bandwidth (MB/s)	Costo Scheda in Euro
Fast- Ethernet	100	12	< 10
Gigabit- Ethernet	100	125	> 10
Myrinet	7	250	> 100
QSWnet	5	360	> 100
InfiniBand	4	750	> 1000

- La **latenza** di trasmissione di un collegamento è il tempo necessario (espresso in millisecondi) impiegato da un **pacchetto Dati** a raggiungere un altro computer o server in rete. Nell'ambito delle reti i fattori che influenzano maggiormente la propagazione del segnale sono il mezzo che trasporta l'informazione e le apparecchiature (p. e. switch o router) che il segnale attraversa nel suo percorso.
- La **bandwidth** è la capacità di trasporto delle informazioni, generalmente espressa in bit per secondo. Maggiore è la larghezza, maggiore è il numero degli utenti che potrà supportare, e maggiori saranno le quantità di informazioni che quegli utenti riusciranno a trasmettere e ricevere.

- **Approccio e sviluppo al problema e al programma**
- Indubbiamente, il primo passo nello sviluppo di un problema scientifico è capire se esso può essere risolto in parallelo.
- **Esempio di problema parallelizzabile.**
- **Calcola la energia potenziale per ognuna delle migliaia di configurazioni indipendenti di una molecola. Eseguiti questi calcoli, trova la configurazione con la energia minima.**
 - Questo è un problema che può essere parallelizzabile. Ognuna delle configurazioni molecolari è determinabile in modo indipendente. **Anche il calcolo della configurazione di energia minima è un problema che può essere reso parallelo.**

- Nuovo esempio di problema:
- **Calcolo delle serie di Fibonacci**
(1,1,2,3,5,8,13,21,.....) usando la formula
- $$F(k + 2) = F(k + 1) + F(k)$$

**E' possibile rendere parallelo
il seguente problema ?????**

- Nuovo esempio di problema:
- **Calcolo delle serie di Fibonacci (1,1,2,3,5,8,13,21,.....) usando la formula**
- **$F(k + 2) = F(k + 1) + F(k)$**

La risposta è NO.

Il calcolo della sequenza di Fibonacci impone calcoli dipendenti piuttosto che indipendenti. Il calcolo del valore **k+2** usa il calcolo dei valori **k+1** e **k**. Questi tre termini non possono essere calcolati in modo indipendente e quindi, non è possibile calcolarli in modalità parallela.

- **Terminologia: Calcolo distribuito**
- Il *calcolo distribuito* (distributed computing) utilizza le risorse di diversi computer collegati in rete tra di loro (solitamente attraverso internet) per risolvere problemi computazionali a larga scala. Semplificando, un'infrastruttura per il calcolo distribuito consiste in un software principale che usa i processori di computer connessi ad una rete (locale o globale) come se fossero processori collocati sulla stessa scheda madre di un calcolatore adibito al calcolo parallelo. Per realizzare ciò è necessario che i computer partecipanti abbiano un software che sfrutta i tempi morti di utilizzo della CPU in maniera da non ridurre le prestazioni delle piattaforme su cui lavorano, elaborando piccole parti di un progetto di ricerca che rimane molto vasto.

- **Terminologia: Grid computing**
- Il grid computing è un *sottoinsieme* molto utilizzato del calcolo distribuito.
- Il termine grid computing (letteralmente, griglia di calcolo) indica un'infrastruttura distribuita per consentire l'utilizzo di risorse di calcolo e di storage, provenienti da *un numero indistinto di calcolatori* (anche e soprattutto di potenza non particolarmente elevata), posti anche a distanze elevate tra di loro, interconnessi da una rete, possibilmente ad alta velocità di trasmissione.

ENEA-GRID

ENEA, la Agenzia nazionale per le nuove tecnologie, l'Energia e lo sviluppo economico sostenibile, ha 12 siti di ricerca, 6 di essi con propri centri di calcolo. Tutte le principali risorse di calcolo dell'Enea sono integrate nella infrastruttura ENEA-GRID che fornisce ai ricercatori ENEA e ai loro collaboratori un facile accesso alle risorse multiplatforma disponibili.

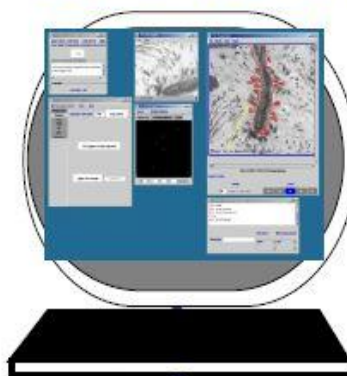


ENEA-GRID Infrastructure

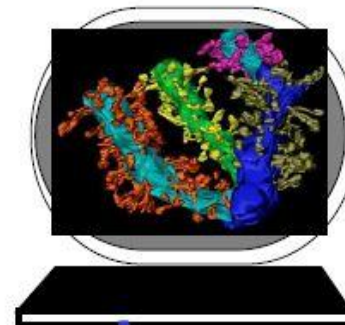
DATA ACQUISITION



DATA ANALYSIS



ADVANCED COMPUTER GRAPHICS



NETWORK



IMAGING INSTRUMENTS



COMPUTATIONAL RESOURCES



MULTI-SCALE DATABASES

HARDWARE

- **Più rilevanti:** sistema HPC CRESCO, situato a Portici (NA)
~ 17.1 Tflops, 300 hosts, ~ 3000 cores, InfiniBand 4xDDR
- **Altri:** ~100 hosts ~650 cpu
 - AIX: IBM SP5 256 cpu (12 p575 1.5GHz, 16 cpu + 1 p595 1.9 Ghz, 64 cpu, 1.5 Tflops); SP4, 96 cpu
 - SGI Altix 350 (IA64) 32 cpu & Onyx4
 - Cray XD1 24 cpu
 - Linux clusters 32/x86_64; Apple cluster; Windows servers....
- **Cluster grafici per rendering 3d remoto + visualizzatore stereo**
- **Strumenti:** Microscopio elettronico, tavola sismica,

SOFTWARE:

- Codici commerciali (Fluent, Ansys, Abaqus,...)
- Codici di ricerca e open source (CPMD, MCNP, OpenFoam,...)
- Strumenti di calcolo (Matlab, IDL,...),
- Software di visualizzazione dati (AVS/Express, Visit,...)

Funzionalità GRID (autenticazione, autorizzazione, scoperta e gestione delle risorse) sono fornite da componenti multiplatforma consolidate (supporti per ENEA-GRID):

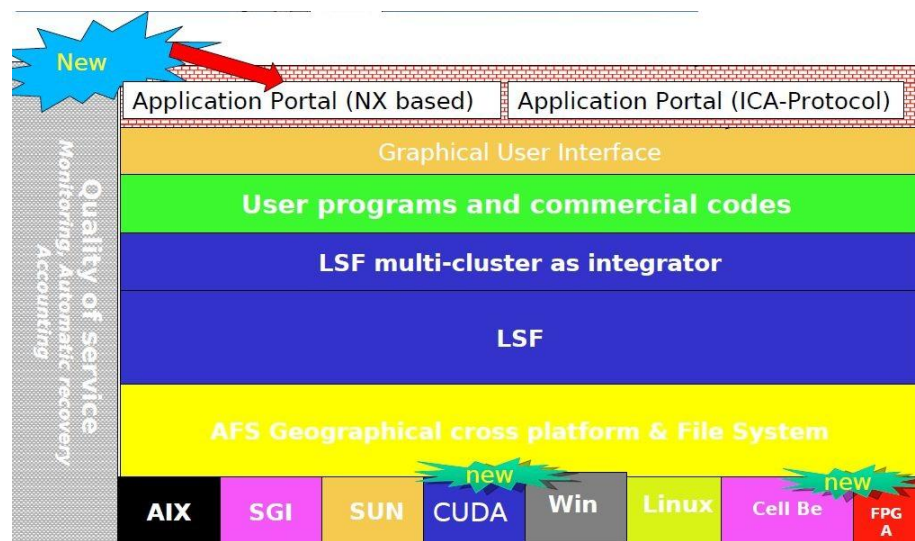
File System distribuito: **OpenAFS/Kerberos 5 integrated**

Gestione di risorse WAN: **LSF Multicluster [www.platform.com]**

Interfacce utente: **Java & Citrix Technologies, now also Freenx**

Queste realizzazioni hanno permesso di integrare in modo affidabile attraverso gli anni (ENEA-GRID ebbe inizio nel 1999) lo stato dell'arte delle risorse di calcolo presenti.

ENEA partecipa in vari **progetti** GRID (Datagrid, EGEE, EGEE-II & III, BEINGRID, PON projects, GRISU..) concentrandosi su **soluzioni di interoperabilità** con gli altri middleware (gLite, Unicore): un **metodo di implementazione di gateway** (Shared Proxy Approach for GRID Objects) è stato sviluppato ed applicato alla interoperabilità con le grid basate su gLite.



- Requirements
 - Computing power: High Performance / High Throughput Computing
 - Direct access to your environment and application
 - Reliable and stable user environment
 - Security and traceability
 - Intellectual property protection
- ENEA-GRID & Research: ENEA Examples
 - Computational Chemistry
 - Nuclear Fusion, plasma stability
 - Climate/Weather/Ocean Simulations
 - Pollutant Atmospheric Diffusion
 - Combustion Simulation
 - Remote 3D Rendering
 -

CLOUD COMPUTING

«Cloud», una delle parole più pronunciate negli ultimi mesi. Letteralmente significa «nuvola». La nuvola ha da sempre rappresentato in modo grafico «Internet». La nuvola, il cloud e la rete Internet sono quindi strettamente collegati.



La Rete, nata nel 1969 con il nome di Arpanet, si è evoluta nel tempo fino al 1991, anno in cui con la definizione del protocollo http e la nascita del world wide web (WWW) prende il nome di Internet.

L'intento fu quello di connettere tra loro più computer con lo scopo di condividere informazioni, documenti, risorse.

I PC, quindi, si possono connettere a Internet, ma nessuno (o pochi) sanno realmente cos'è, cosa c'è dietro, come avviene la connessione e quali sono i servizi che consentono di farla.

Non so dove è pubblicato il sito Web che visito e non so neanche come è stato creato.

Sto quindi usufruendo di un servizio che una piattaforma, qualunque essa sia, mi sta erogando. E' qui che nasce il concetto di NUVOLA.

Cloud Computing.

Un insieme di risorse hardware e software che forniscono servizi su richiesta attraverso la rete Internet.

Attenti a non confondere il Cloud con Internet.

Il cloud è sostanzialmente un servizio fornito da un gestore di terze parti con la formula «on demand».

L'utente paga soltanto per quello che consuma, in termini di banda, di risorse, di tempo di utilizzo, di numero di transazione, di utilizzo di memoria.

Questo modello permette di creare soluzioni scalabili, performanti e affidabili anche con basso investimento iniziale, senza cioè dover acquistare l'hardware necessario e i servizi connessi: consumo di elettricità, disponibilità di una connessione veloce e permanente a Internet.

Caratteristica fondamentale e imprescindibile del Cloud Computing è quella di fornire una infrastruttura di servizi assolutamente **affidabile** e **scalabile**. E' infatti possibile, a seconda delle necessità, aumentare o diminuire i nodi in uso, oppure cambiare le caratteristiche dei server potenziando CPU e memoria.

Tutto questo è possibile perché l'infrastruttura è basata su un sistema di virtualizzazione delle risorse. La piattaforma di calcolo viene eseguita su **macchine virtuali** opportunamente dimensionate e in grado di fornire i servizi richiesti. Ogni volta che viene richiesto di scalare, utilizzando un nuovo nodo, la piattaforma si preoccupa di replicare l'applicazione sul nuovo nodo e renderlo quindi disponibile.



Dunque

Grid Computing:

Offre un accesso controllato alle risorse condivise, pur garantendo all'utente la visibilità di un unico sistema di calcolo al quale sottomettere i propri programmi (job).

Cloud Computing:

L'implementazione effettiva delle risorse non è definita, né le sue caratteristiche sono note all'utilizzatore.

Per semplificare: anziché sul computer, i software vengono installati direttamente sulla rete, appunto in una «nuvola». I dati che fino ad oggi venivano salvati sui pc sono decentrati su vari server: giganteschi archivi digitali a cui l'utente può accedere grazie al browser e alle applicazioni. In pratica , invece di archiviare la nostra musica e i nostri documenti sul computer di casa, li depositeremo su Internet e potremo consultarli con diversi dispositivi.



Vantaggi.

Con la tecnologia Cloud si possono creare degli archivi raggiungibili in qualunque momento. Dalle fotografie delle vacanze ai documenti aziendali e alle proprie cartelle sanitarie. Tutto a portata di smartphone e di iPad.

Rischi.

In caso di black-out o incidenti al server, i servizi e i dati non saranno raggiungibili.

I rischi più seri riguardano invece la privacy e la censura. I dati, le informazioni e il potere da essi derivanti rischiano di venire concentrati nelle mani di pochissimi grandi gruppi, in grado di bloccare il cyberspazio per le voci scomode.

GRAZIE PER L'ATTENZIONE