

# **Uma breve história da supercomputação**

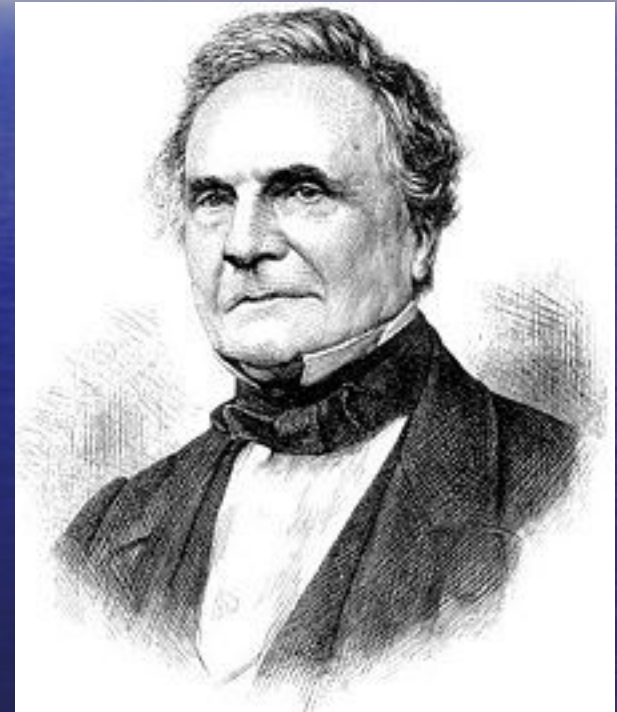
**Stephan Stephany - LAC/INPE  
seminário CAP – 03/10/2012**

**CAP/Computação Aplicada - INPE**

# A gênese dos computadores:

Charles Babbage (1791-1871)

- **Difference engine**, uma calculadora mecânica 13 t
- **Analytical engine**, seria o 1º. computador mecânico, nunca concluído, e que teve **Ada Lovelace** (1815-1852), matemática, como a 1ª. programadora história

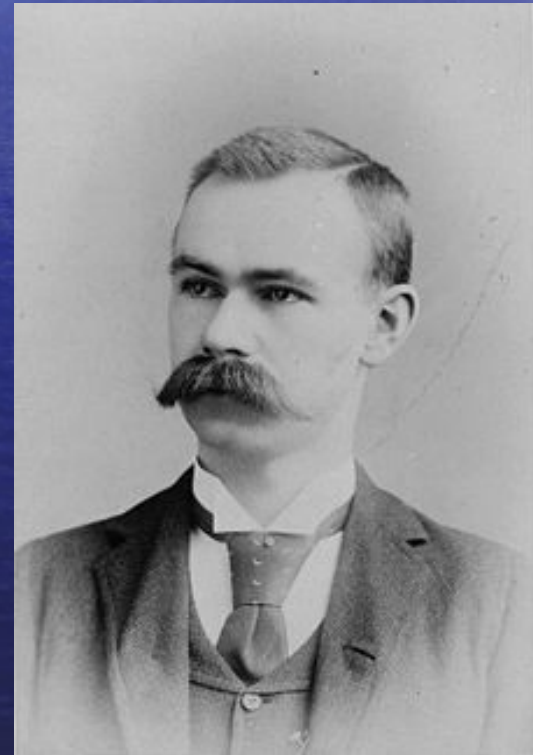


# A gênese dos computadores:

## Herman Hollerith (1860-1929)

**Máquina tabuladora de dados com cartões perfurados para o censo dos EUA de 1890).**

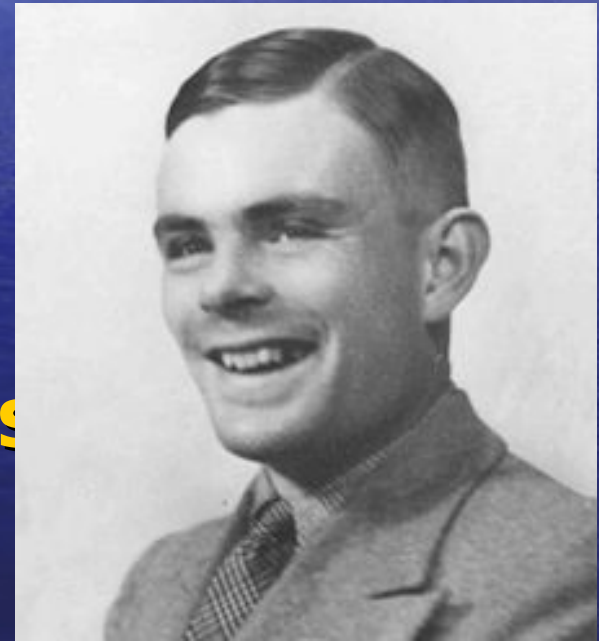
**Sua empresa deu origem em 1924 à IBM (International Business Machines)**



# **A gênese dos computadores:**

## **Alan Turing (1912-1954)**

**Conceito de máquina de Turing (1936), uma máquina de estados finitos capaz de executar um algoritmo.**



# A gênese dos computadores: primeiros computadores

- **Z3 de Konrad Zuse (Alemanha, 1941)**
- **Colossus de Tommy Flowers (Inglaterra, 1944)**
- **Harvard Mark I de Howard Aiken (EUA, 1944)**
- **ENIAC de Eckert e Mauchly (EUA, 1944)**
- **Atanasoff-Berry Computer ABC (EUA, 1941)**

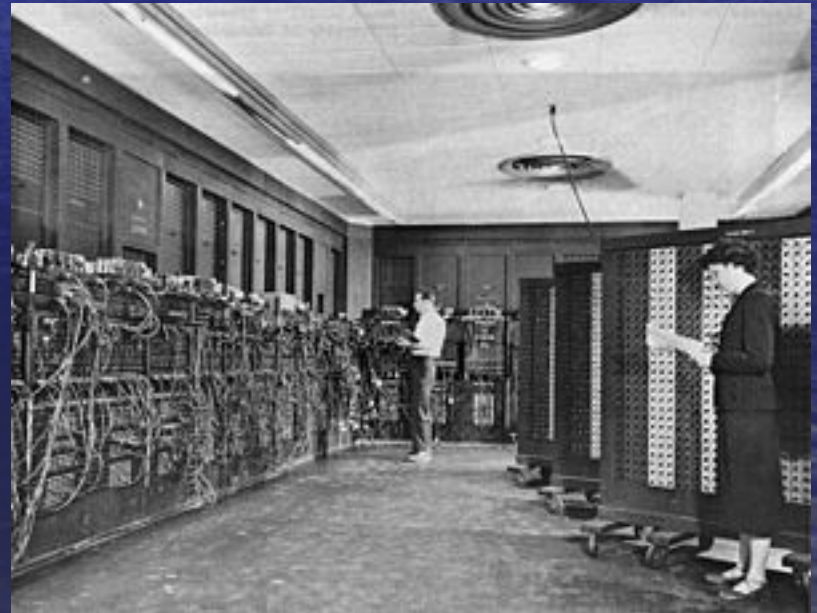
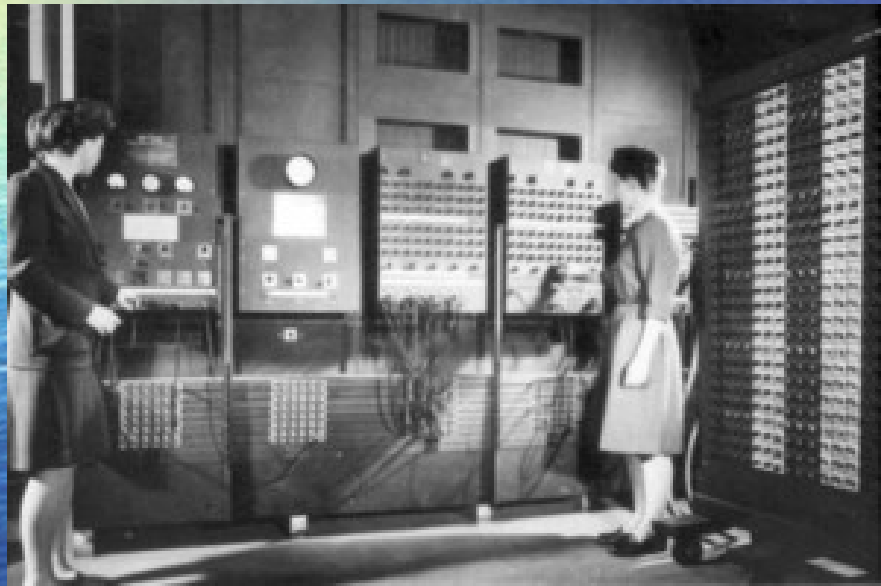
**Atanasoff-Berry Computer ABC** (EUA, 1941) foi o 1o. comp. eletrônico digital, binário, “a válvula”, não Turing-completo, para sistemas de eqs. lineares.

**Z3 de Konrad Zuse** (Alemanha, 1941) foi o 1o. computador digital de propósito geral, binário, eletromecânico, Turing-completo, programável por cartão perfurado.

**Colossus de Tommy Flowers** (Inglaterra, 1944) comp. eletrônico digital, binário, “a válvula”, não Turing-completo, para criptoanálise.

**Harvard Mark I de Howard Aiken** (EUA, 1944) comp. programável por cartão perfurado, decimal, “a relê”.

# A gênese dos computadores: ENIAC - Electronic Numerical Integrator And Computer (1943-1955)



# **A gênese dos computadores:**

## **ENIAC - Electronic Numerical Integrator And Computer (1944)**

- computador eletrônico digital, decimal, “a válvula”, Turing-completo, de propósito geral, programável “à mão”.**
- John Presper Eckert (1919-95) & John Mauchly (1907-80)**
- John von Neumann (1903-57)**



# A gênese dos computadores:

## Sperry Rand UNIVAC

- John Presper Eckert & John Mauchly formaram a Eckert & Mauchly Computer Co.
- Grupo de pesquisadores da US Navy formou a Engineering Research Associates (ERA)
- Ambas foram compradas pela Sperry Rand



# A gênese dos computadores:

## Sperry Rand UNIVAC

**-Sperry Co. comprou em 1955 a Remington Rand e surgiu a Sperry Rand. Em 1978 vendeu divisões e voltou a ser Sperry Co. que em 1986 fundiu-se com a Burroughs para formar a UNISYS.**

**-UNIVAC I (*UNIVersal Automatic Computer I*) foi o primeiro computador vendido comercialmente nos EUA (1951) com 46 sistemas vendidos, sendo usado no censo de 1951 e também na eleição presidencial de 1952.**

# A gênese dos computadores:

## CDC - Control Data Co.

- **Snow White (IBM) & 7 dwarfs:** Burroughs, NCR, Control Data Corporation, General Electric, RCA, Honeywell.
- **IBM & “the bunch”** (Burroughs, Univac, NCR, Control Data, Honeywell).
- **Em 1957 um grupo da antiga ERA saiu da UNIVAC/Sperry-Rand para formar a Control Data Co. incluindo Seymour Cray.**

# A gênese dos computadores:

## CDC - Control Data Co.

-CDC-6600 (1964) - 1o. supercomputador da história, 500 Kflop/s - 1 Mflop/s (10 x mais rápido que concorrentes como o IBM-360) em projeto liderado por Seymour Cray.

**Thomas J. Watson Jr. (IBM):** *How is it that this tiny company of 34 people can be beating us when we have thousands of people?* **Cray:** *You just answered your own question.*

-CDC-7600 (1969) - 2 Mflop/s, mas pouco confiável (problemas h/w), pipelining.

# Supercomputadores:

## Seymour Cray (1925-1996)

Deixou a CDC em 1972 para fundar a **Cray Research Inc. (CRI)** que projetou o CRAY-1 (1976, 250 Mflop/s), sucesso comercial

CRAY X-MP (1982) 4 x 200 Mflop/s

CRAY-2 (1985) 1.9 Gflop/s

CRAY Y-MP (1988) 8x333 Mflop/s



Cray Computer Co. (CCC, 1989) faliu (1995).  
CRI fundiu-se c/ Silicon Graphics (SGI) em 1986. Atual Cray Inc. surgiu em 2000.

# **Supercomputadores:**

## **CRAY-1 (1976)**

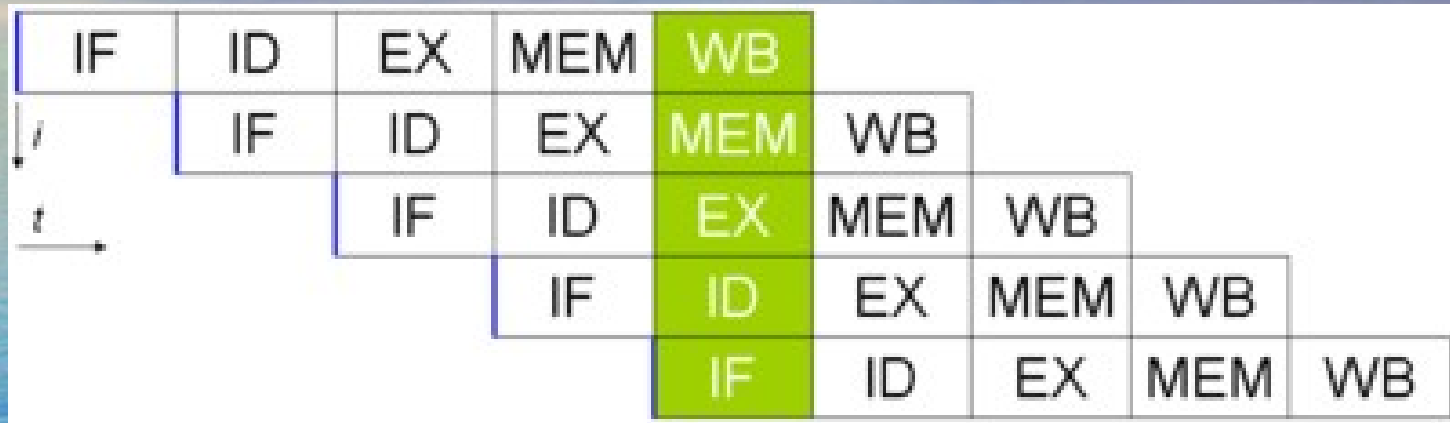
- **1 processador vetorial, uso pioneiro de circuitos integrados, pipelining, registradores escalares e vetoriais, clock de 12.5 ns (80 MHz), 24-bit memory 8 MB em 16 bancos com ciclo 50 ns. Desempenho de 160 mip/s (2x80 MHz) ou 136 Mflop/s, chegando a 250 Mflop/s com instruções vetoriais, refrigerado a freon.**
- **sucesso comercial: 80 unidades a US\$5-8 para governo EUA, etc.**

# Supercomputadores:

## CRAY-1 (1976) & CRAY-2 (1985)



# A busca por desempenho: **PIPELINING**



1 CPU 1 GHz e 2 pipelines FP = 2 Gflop/s?

-Fluxo de instruções independentes!

-todos estágios duração 1 ns (max)

-Ex: AMD Opteron 12/17 estág. INT/FP

# A busca por desempenho: **RISC x CISC**

- **Hardware** -> <- **software**
- **CISC** (complex instruct. set computing)
- **Microprogramação**
- **RISC** (reduced instruct. set computing)  
“simples e poucas”, melhor pipelining,  
instr. exclusivas acesso memória
- máquinas atuais: “mistura” CISC/RISC

# A busca por desempenho: **VETORIZAÇÃO**

- **processadores vetoriais: instruções vetoriais, registradores vetoriais e pipelining eficiente - em extinção?**
- **vetorização implementada em máquinas escalares (PCs) por instruções MMX/SSE e suportada por compiladores; vetorização nos SPE's dos cell processors.**
- **uso de FPGA's**

# **A busca por desempenho:** **OTIMIZAÇÃO DA MEMÓRIA**

- clock dos processadores é sempre mais rápido que o da memória...
- uso de hierarquia de memória (muitos níveis de cache), memória virtual, etc.
- Arquiteturas para evitar contenção p/ acesso à memória entre procs./cores.
- cache hit > 95% !!!!
- memória abundante (64-bits) e barata

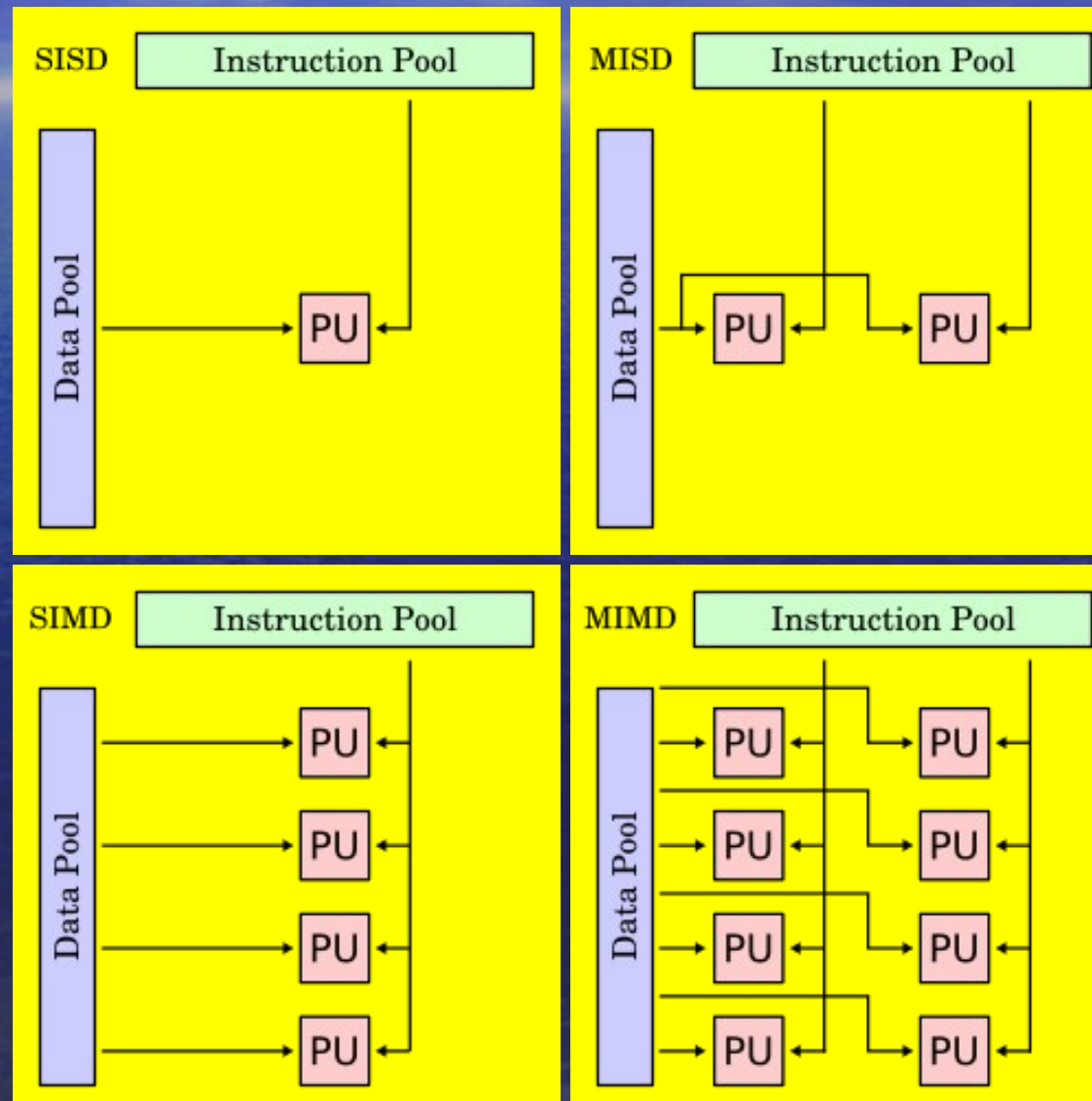
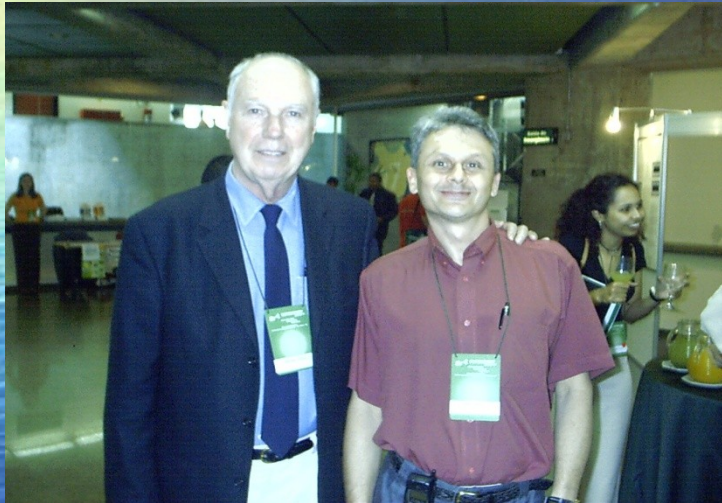
# A busca por desempenho: **PARALELISMO**

- Instruction level parallelism - explora técnicas para executar várias instruções concorrentemente - depende do H/W do processador e compilador otimizante (fluxo contínuo instruções independentes)
- multi-threading
- Multiprocessamento exige programação paralela, bibliotecas (MPI, OpenMP, etc.)

# A busca por desempenho: **PARALELISMO**

- escalabilidade - vai rodar mais rápido (speed-up) com mais processadores?
- granularidade, custos de computação e de comunicação.
- balanceamento de carga
- gargalos sequenciais, lei de Amdahl
- novos algoritmos paralelos, algoritmos tolerantes a falhas, etc.

# A busca por desempenho: **TAXONOMIA DE FLYNN**



# A busca por desempenho: **BENCHMARKS**

- benchmarks - o que são?
- desempenho continuado, de pico, nominal...
- resultados tendenciosos...(Gflop/s?)
- benchmarks: [www.spec.org](http://www.spec.org)
- lista TOP-500: [www.top500.org](http://www.top500.org)
- e o seu software?

# Quem precisa de supercomputadores?

- **previsão do tempo, modelagem ambiental, mecânica dos fluídos computacional, bombas atômicas, sismologia, etc.**
- **indústria do petróleo (Petrobrás)**
- **indústria entretenimento, Internet**
- **decriptografia, contraespionagem, etc.**
- **serviços públicos (correios, receita, etc.)**

# Quem precisa de supercomputadores?

- 5 centros de supercomputação do National Science Foundation (EUA), como o NCSA da Univ. Illinois at Urbana-Champaign.
- supercomputadores dos laboratórios nacionais EUA como o ASCI Blue Mountain de Los Alamos.
- 7 CENAPAD's - Sistema Nacional Processamento de Alto Desempenho (CPTEC/INPE, COPPE/UFRJ, UFMG, UNICAMP, UFRGS, UFPE, LNCC)

# Supercomputadores:

## Deep Blue IBM x Garry Kasparov

- para jogar xadrez x campeão mundial
- venceu torneio em 1997
- 30 procs. RISC/6000 & h/w específico (xadrez)
- 11.38 Gflop/s



# Supercomputadores:

## Earth Simulator (Yokohama, Japão)

-O mais rápido **2002-2004** para executar modelos climáticos globais, de aquecimento global e de geofísica, modelando atmosfera e oceanos.

arquitetura do NEC SX-6: 640 nós x 8 procs.  
Vetoriais = 5.120 ou 56 Gflop/s por nó ou 36 Tflop/s totais, 16 GB memória, 700 TB disco e 6.400 kilowatt de consumo de eletricidade !!! (1.500 chuveiros elétricos)

# Supercomputadores:

## **Cyclops64** da família **IBM Blue Gene**

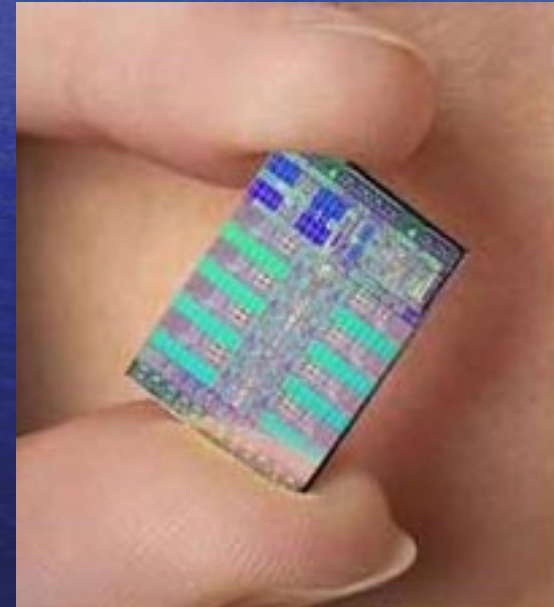
- cada chip/cell (80 Gflop/s nominais):  
arquitetura celular, clock de 500 MHz e 80 processadores ponto-flutuante ligados por crossbar switch e acesso a memória intercalada explorando paralelismo a nível de threads.
- full system 13.824 chips em 96 racks num total de 1.105.920 processadores e 1.1 petaflop/s nominais (1 petaflop =  $10^{15}$ )
- 1 kiloflop =  $10^3$       1 gigaflop =  $10^9$
- 1 megaflop =  $10^6$       1 teraflop =  $10^{12}$

# CELL PROCESSOR

-Arquitetura desenvolvida p/  
Sony/Toshiba/IBM adequada  
para multimídia (Play Station  
3) ou processamento vetorial  
e acesso eficiente à memória.

-1 PPE (Power Processing  
Element) & 8 SPE's co-  
processadores vetoriais  
integrados num chip.

teste com clock 3.2 GHz  
obteve 100 Gflop/s !!!



# COMPUTAÇÃO VERDE

**Diminuição consumo energia (“baixar” clock, usar multi-cores, etc.), reciclagem, etc. Atualmente pode-se estimar 20 W/Gflop/s (apenas processamento)**

- Intel Pentium 100 MHz ( 10 W )
- Intel Pentium IV 3.6 GHz ( 115 W )
- Intel Itanium-2 1.66 GHz ( 122 W )
- Intel Xeon 3.2 GHz ( 110 W )
- Intel Xeon 3.5 GHz dualcore ( 150 W )
- Intel Core Duo 2.16 GHz ( 31 W )
- Intel Pentium M 2.16 GHz ( 27 W )
- AMD Turion-64 2.4 GHz ( 35 W )

# Supercomputadores:

## INPE – CPTEC

**-NEC SX-3 (1994)** - apenas 1 nó, 1 proc. vetorial  
ciclo de 2.5 ns (400 MHz) com 8 pipelines, 3.2  
Gflop/s, refrig. água., 1º. supercomp. do Brasil.

**-NEC SX-4 (1998)** - 1 nó com 8 proc. vetoriais,  
com 2 Gflop/s por proc. ou 16 Gflop/s totais,  
memória de 8 GB, 256 GB disco.



# Supercomputadores:

## INPE – CPTEC

**NEC SX-6 (2002/2004)** - 12 nós x [8 proc. vetoriais] (96), memória 768 GB, 16 TB disco, 50 km cabos, peso 13.5 t, 8 Gflop/s por proc. ou 768 Gflop/s.

**UNA-1 NEC/Sun (2007)** - 275 nós escalares 2 x dual-core Opteron 64-bits 2.6 GHz total 1.100 procs. 5.2 Gflop/s por core ou 5.72 Tflop/s e rede Infiniband.

# Supercomputadores: TUPÃ (2010) - **INPE** – **CPTEC**

**CRAY XE-6** - 65 nós “auxiliares” + 1304 nós  
processamento 2 x 12-core AMD Opteron  
2.1 GHz memória 32 GB/nó, Cray Gemini  
Interconnect, 258 Tflop/s ou 205 (Linpack)



# Ranking Top 500 - mundial

[www.top500.org](http://www.top500.org)

| Rank | Site                                                                  | System                                                                                                        | Cores   | Rmax<br>(TFlop/s) | Rpeak<br>(TFlop/s) | Power<br>(kW) |
|------|-----------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|---------|-------------------|--------------------|---------------|
| 1    | DOE/NNSA/LLNL<br>United States                                        | <b>Sequoia</b> - BlueGene/Q, Power BQC 16C<br>1.60 GHz, Custom<br>IBM                                         | 1572864 | 16324.8           | 20132.7            | 7890          |
| 2    | RIKEN Advanced Institute for<br>Computational Science (AICS)<br>Japan | <b>K computer</b> , SPARC64 VIIIfx 2.0GHz, Tofu<br>interconnect<br>Fujitsu                                    | 705024  | 10510.0           | 11280.4            | 12659.9       |
| 3    | DOE/SC/Argonne National Laboratory<br>United States                   | <b>Mira</b> - BlueGene/Q, Power BQC 16C<br>1.60GHz, Custom<br>IBM                                             | 786432  | 8162.4            | 10066.3            | 3945          |
| 4    | Leibniz Rechenzentrum<br>Germany                                      | <b>SuperMUC</b> - iDataPlex DX360M4, Xeon<br>E5-2680 8C 2.70GHz, Infiniband FDR<br>IBM                        | 147456  | 2897.0            | 3185.1             | 3422.7        |
| 5    | National Supercomputing Center in Tianjin<br>China                    | <b>Tianhe-1A</b> - NUDT YH MPP, Xeon X5670<br>6C 2.93 GHz, NVIDIA 2050<br>NUDT                                | 186368  | 2566.0            | 4701.0             | 4040          |
| 6    | DOE/SC/Oak Ridge National Laboratory<br>United States                 | <b>Jaguar</b> - Cray XK6, Opteron 6274 16C<br>2.200GHz, Cray Gemini interconnect,<br>NVIDIA 2090<br>Cray Inc. | 298592  | 1941.0            | 2627.6             | 5142          |

# Ranking Top 500 - Brasil

[www.top500.org](http://www.top500.org)

(Tupã era 29º em Nov/2010)

| Rank | Site                                                      | System                                                                                       | Cores | Rmax<br>(TFlop/s) | Rpeak<br>(TFlop/s) | Power<br>(kW) |
|------|-----------------------------------------------------------|----------------------------------------------------------------------------------------------|-------|-------------------|--------------------|---------------|
| 68   | PETROBRAS<br>Brazil                                       | Grifo04 - Itautec Cluster, Xeon X5670 6C<br>2.930GHz, Infiniband QDR, NVIDIA 2050<br>Itautec | 17408 | 251.5             | 563.4              | 365.5         |
| 79   | INPE (National Institute for Space<br>Research)<br>Brazil | Tup - Cray XE6, Opteron 6172 12C 2.10GHz,<br>Cray Gemini interconnect<br>Cray Inc.           | 31104 | 214.2             | 261.3              |               |
| 457  | NACAD/COPPE/UFRJ<br>Brazil                                | Galileu - Sun Blade x6048, Xeon X5560 2.8<br>Ghz, Infiniband QDR<br>Oracle                   | 6464  | 64.6              | 72.4               | 430           |

# **Recursos supercomputação acessíveis para projetos**

**CCST/INPE**

**CPTEC/INPE**

**FAPESP**

**INCT**

**Rede Clima**

# **Processamento de Alto Desempenho na CAP**

**curso PAD - profs. Airam J. Preto,  
Stephan Stephany (prof. Celso Luiz  
Mendes nos EUA), formação de  
competências na área.**

**mestrados e doutorados em PAD, com  
ênfase na paralelização de aplicativos.**

**-**

**3 servidoras paralelas 2 x Xeon 2.93 GHz  
quadcore memória 32 GB 6 TB disco**

**demanda por PAD (CPTEC, Embraer,  
Petrobrás, institutos de pesquisa, etc.**

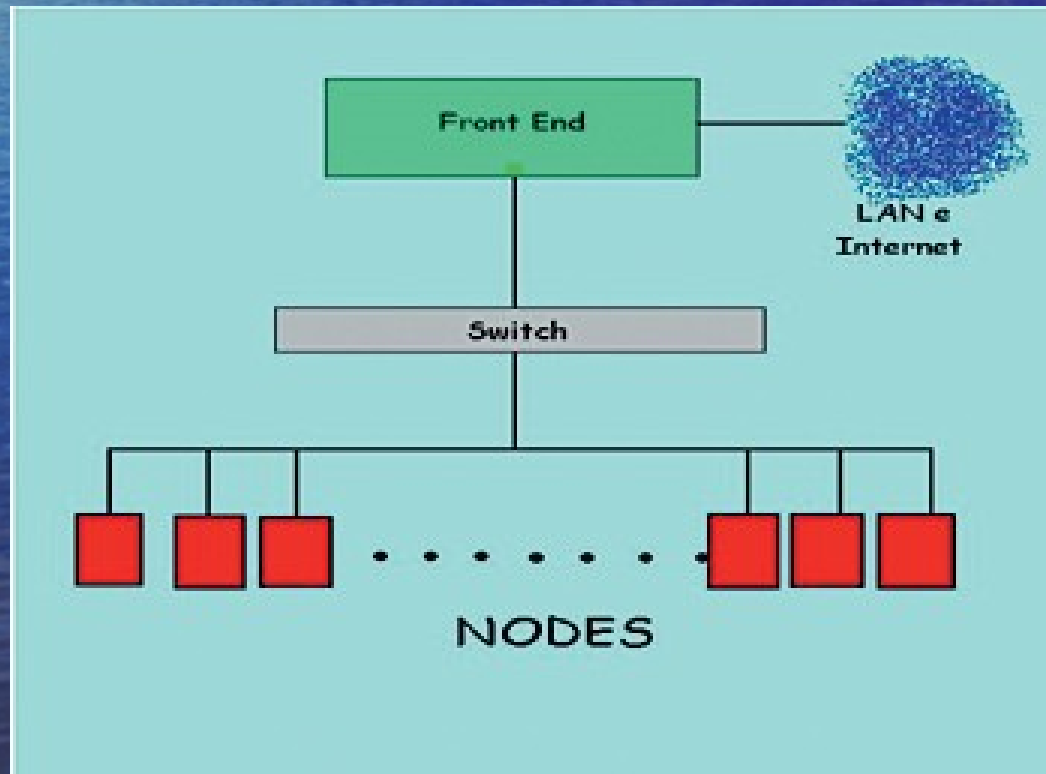
# Cluster Beowulf LAC 2001 (FAPESP)

- desativado fins 2005 R\$40.000 + suite compiladores com 17 nós monoprocc. single-core, AMD 1.33 GHz (upgrade p/ 1.66 GHz), 512 MB/nó, switch Fast Ethernet (100 Mbits/s).

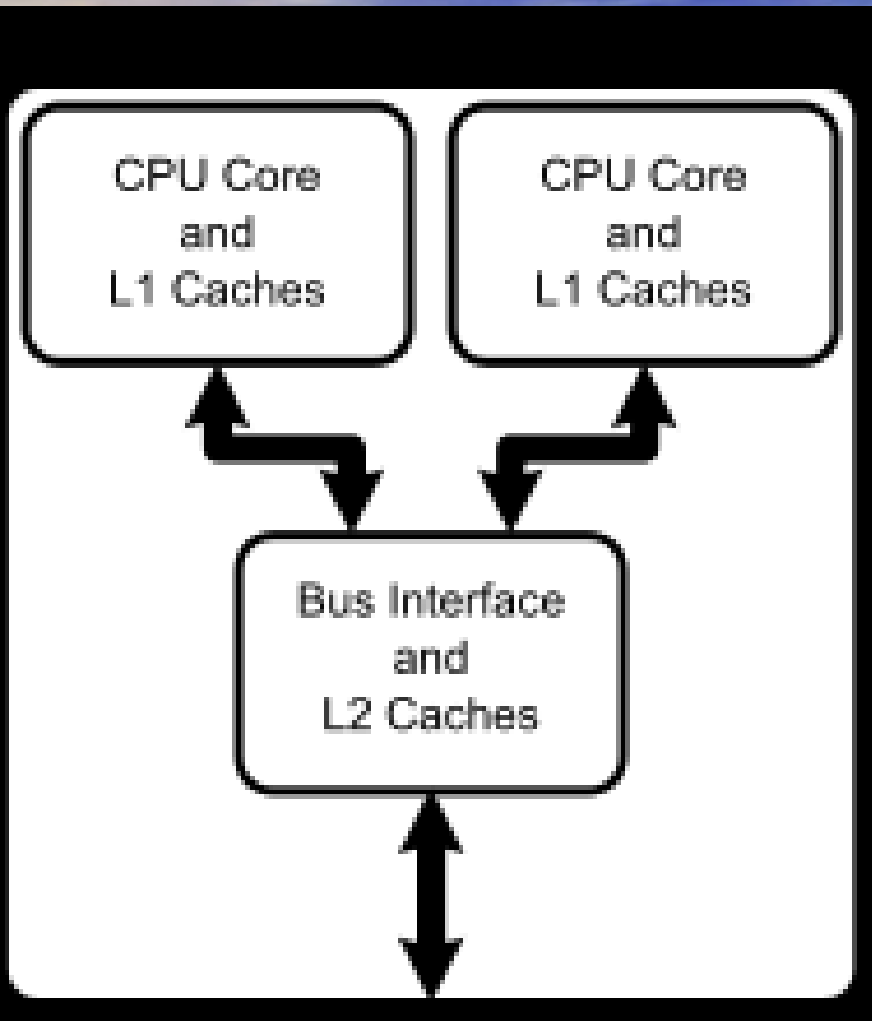


# MPI (Message Passing Interface)

Biblioteca de comunicação de troca mensagens ([www.mcs.anl.gov/mpi/](http://www.mcs.anl.gov/mpi/)); pode ser usada em arquiteturas memória compartilhada ou distribuída e com F77/F90 /C/C++



# Arquiteturas multi-core



- dual-core, quad-core, hexacore...

-pipelining, superscalar, multithreading

-baixo consumo energia (clock baixo), compactas

# Programas para arquiteturas multi-core

- ❖ **Eficiência = escolha algoritmo conveniente + bom compilador (uso dos pipelines, acesso à memória, ou seja, executável otimizado) + paralelização adequada (granularidade, comunicação) + hardware (memória, rede, etc.)**
- ❖ **não adianta paralelizar programa ineficiente...**
- ❖ **“Memory wall” ocorre pois o cache L2/L3 e a memória são compartilhados**

# Multithreading

- **surgiu em processadores single-core com múltiplos pipelines provendo mais um nível de paralelização;**
- **sistema operacional provê escalonamento de threads;**

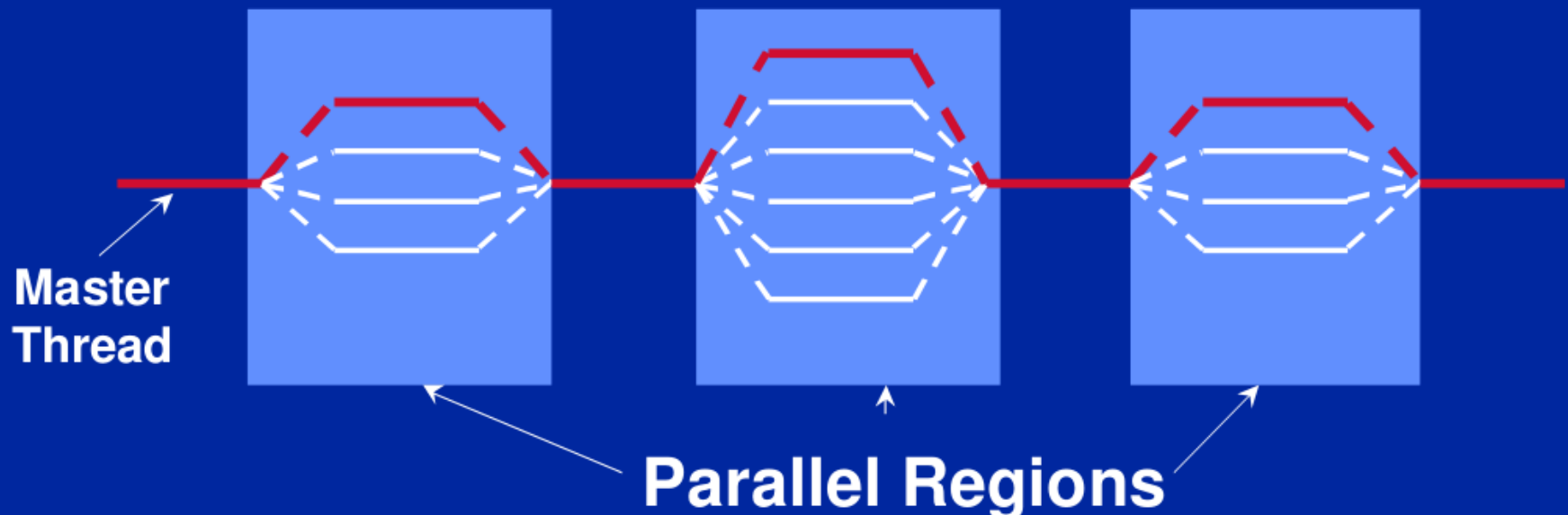
## OpenMP (Multi-Processing)

- **baseado em threads (*[www.openMP.org](http://www.openMP.org)*);**
- **diretivas de paralelização baseada em threads para F77/F90 /C/C++ e execução em arquiteturas de memória compartilhada;**

# OpenMP: Programming Model

## Fork-Join Parallelism:

- ◆ Master thread spawns a team of threads as needed.
- ◆ Parallelism is added incrementally: i.e. the sequential program evolves into a parallel program.



# OpenMP:

## How is OpenMP typically used?

- OpenMP is usually used to parallelize loops:
  - Find your most time consuming loops.
  - Split them up between threads.

Split-up this loop between multiple threads

```
void main()
{
    double Res[1000];

    for(int i=0;i<1000;i++) {
        do_huge_comp(Res[i]);
    }
}
```

Sequential Program



```
void main()
{
    double Res[1000];
    #pragma omp parallel for
    for(int i=0;i<1000;i++) {
        do_huge_comp(Res[i]);
    }
}
```

Parallel Program