

Introdução às Arquiteturas Paralelas

Luís Fabrício W. Góes (Cabra)

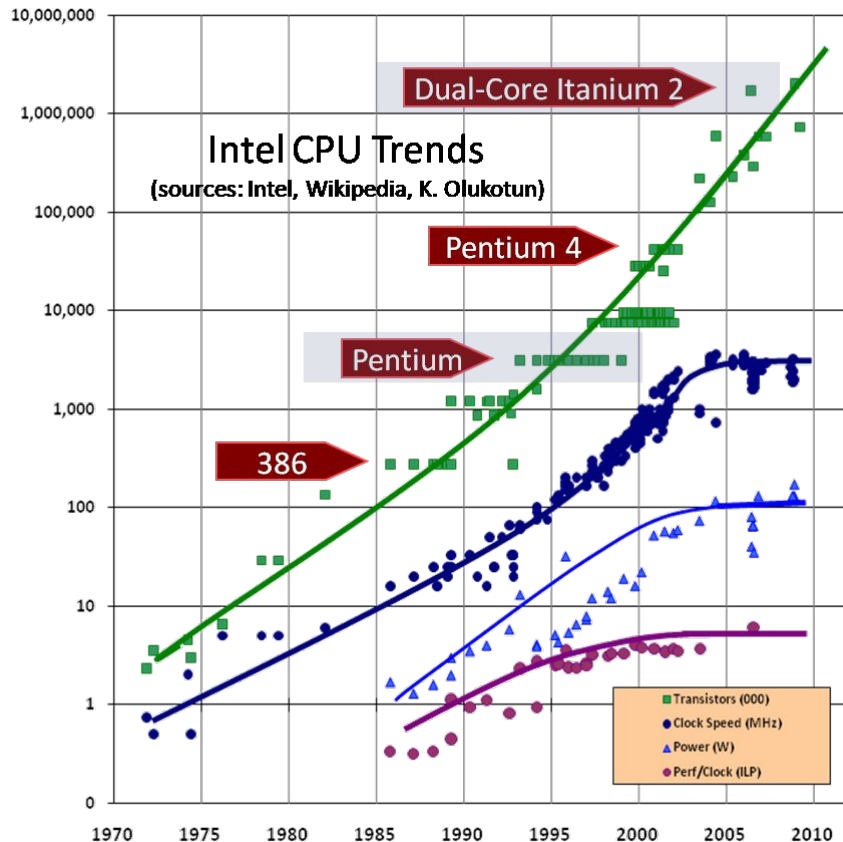
lfgwoes@pucminas.br

“Supercomputador Santos Dumont localizado em Petrópolis, RJ, no Laboratório Nacional de Computação Científica (LNCC). É o maior supercomputador do Brasil. Hoje ele ocupa a posição 471 no TOP500, com 10 mil núcleos e desempenho de até 456 Teraflops.”

Fonte: Incc.br

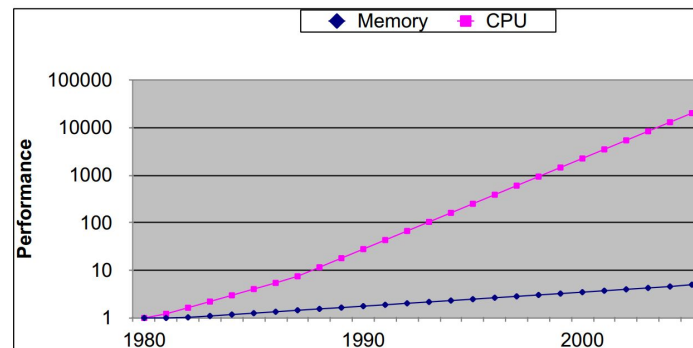
Lei de Moore

- Número de transistores em um chip dobra a cada 18 meses
- O que fazer com estes transistores?
 - Aumentar estágios do pipeline
 - Aumentar frequência
 - Aumentar tamanho da memória cache
- Em 2003 estas práticas pararam de melhorar o desempenho dos processadores de um único núcleo.



Os três muros! Por que multicores?

- Power Wall
 - Fim da lei de Dennard em 2006, onde a relação entre tamanho dos transistores e consumo de energia era constante.
- Instruction Level Parallelism (ILP) Wall
 - A dependência de dados entre instruções limita o tamanho dos pipelines.
- Memory Wall
 - Busca de dados na memória retarda (delay) o processamento em mil ciclos.



\sum muro_i + lei de Moore \rightarrow Multicore

O que são multicores?

- Multicores são processadores com múltiplos núcleos em um mesmo chip, que cooperam entre si para solução de problemas computacionais de forma rápida.
- Qual o problema deles?
 - Paralelismo passa a ser responsabilidade do programador, não mais do compilador ou HW.
- Mas isso é tão ruim assim? Com o que o programador tem que se preocupar?
 - Criação e sincronização de threads
 - Balanceamento de carga
 - Afinidade de memória
 - Deadlocks
 -

Multicores atuais

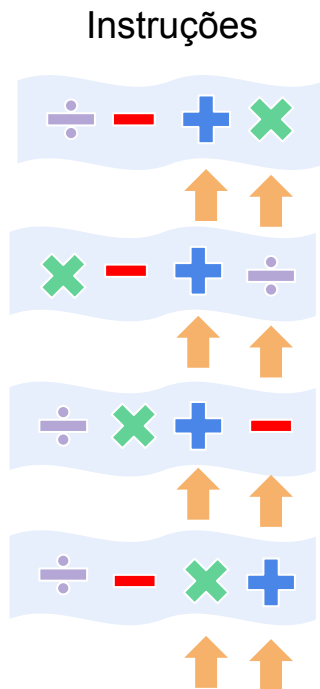


Processador	Intel® Core™ i9-7900X X-series Processor	AMD Ryzen™ Threadripper 1950X	IBM POWER 9
Número de Núcleos	10	16	24
Número de Threads	20	32	96
Frequência Base	3.3 GHz	3.4 GHz	4 GHz
Tamanho da Cache L3	13.75 MB	32 MB	120 MB

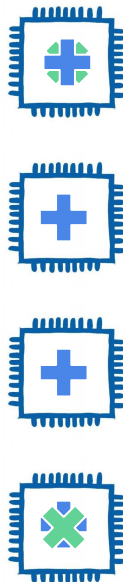
Somente multicores são arquiteturas paralelas?

- Não. Existem diversas arquiteturas paralelas, muito antes dos multicores, divididas pela **taxonomia de Flynn** de acordo com o fluxo de instruções e dados
 - MIMD (multiple instruction multiple data)
 - SIMD (single instruction multiple data)
 - MISD - não faz sentido
 - SISD é o processador sequencial
- As arquiteturas MIMD ainda podem ser divididas em
 - Memória compartilhada
 - Memória distribuída

Multiple Instruction Multiple Data MIMD



Núcleos



Multicores
(memória compartilhada)

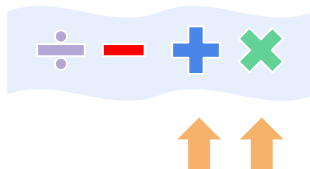


Aglomerados de Computadores (clusters)
(memória distribuída)

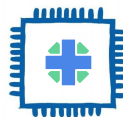
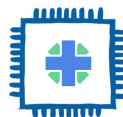
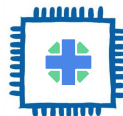
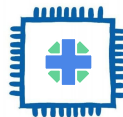


Single Instruction Multiple Data SIMD

Instruções



Núcleos



Máquinas vetoriais

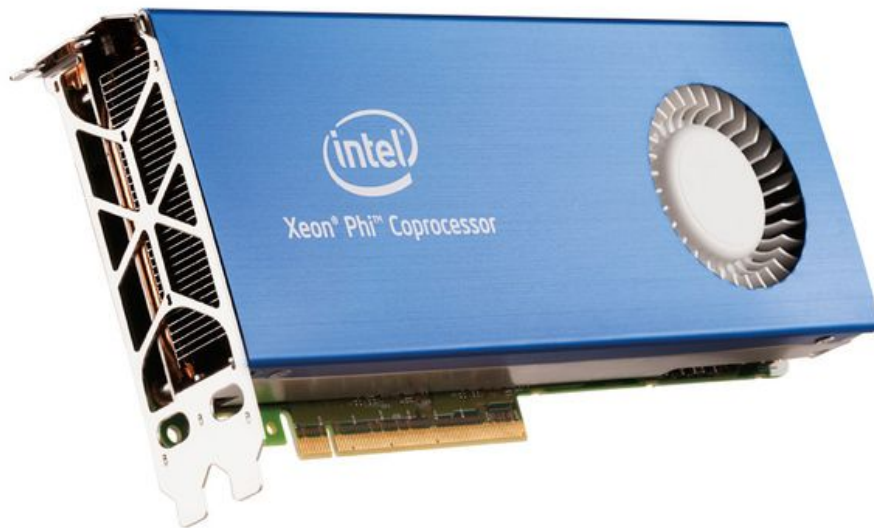


Graphics Processing Unit (GPU)



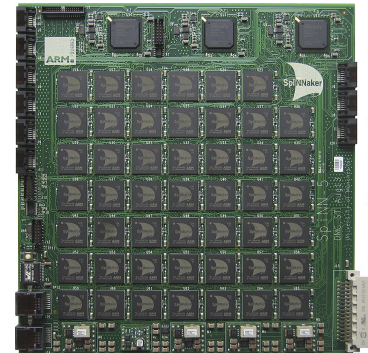
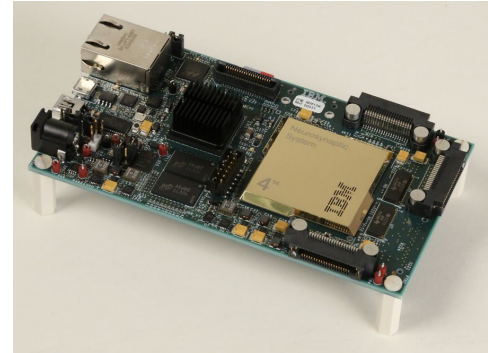
E ainda existem os manycores ...

- O Xeon Phi possui em torno de 60 núcleos com ULAs de 512 bits que permitem vetorização (instruções SIMD).
- Na prática, a maioria das arquiteturas hoje são uma junção de MIMD e SIMD, inclusive os multicores.

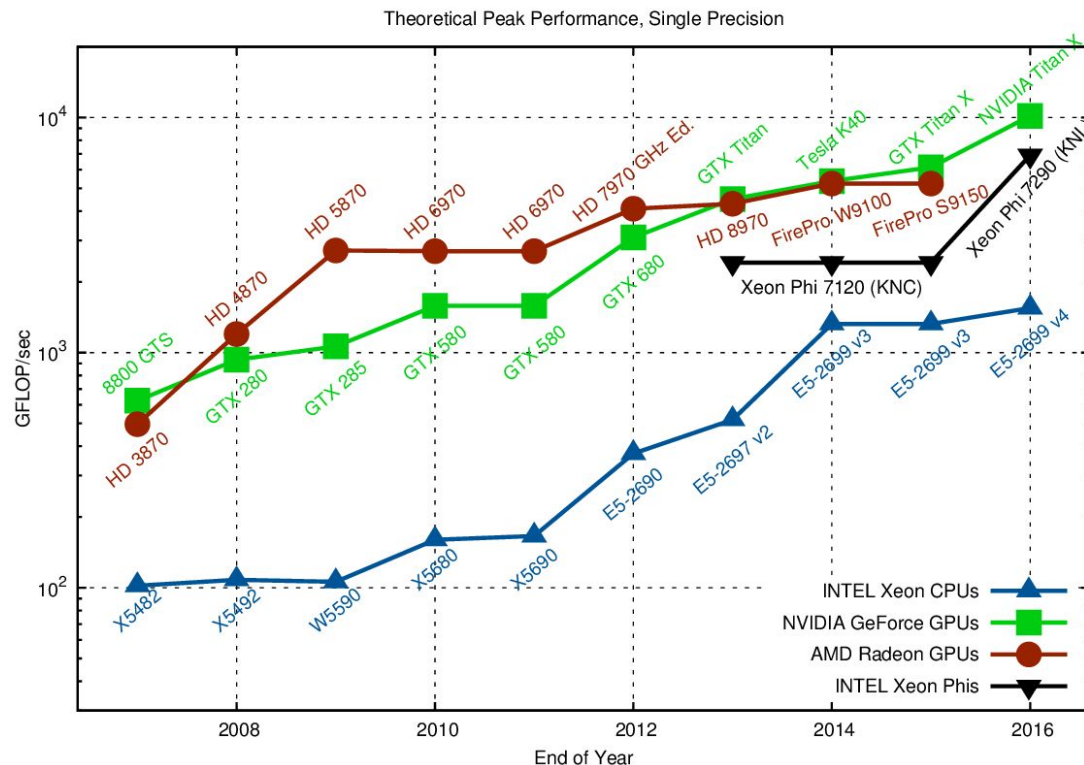


E as arquiteturas neuromórficas ...

- O IBM TrueNorth possui 4096 processadores sinápticos que simulam mais de um milhão de neurônios com baixo custo de energia.
- O SpiNNaker utiliza 1 milhão de processadores ARM para simular o cérebro humano.



Desempenho Multicores, Manycores e GPUs



Fonte: <https://www.karlrupp.net/2013/06/cpu-gpu-and-mic-hardware-characteristics-over-time/>

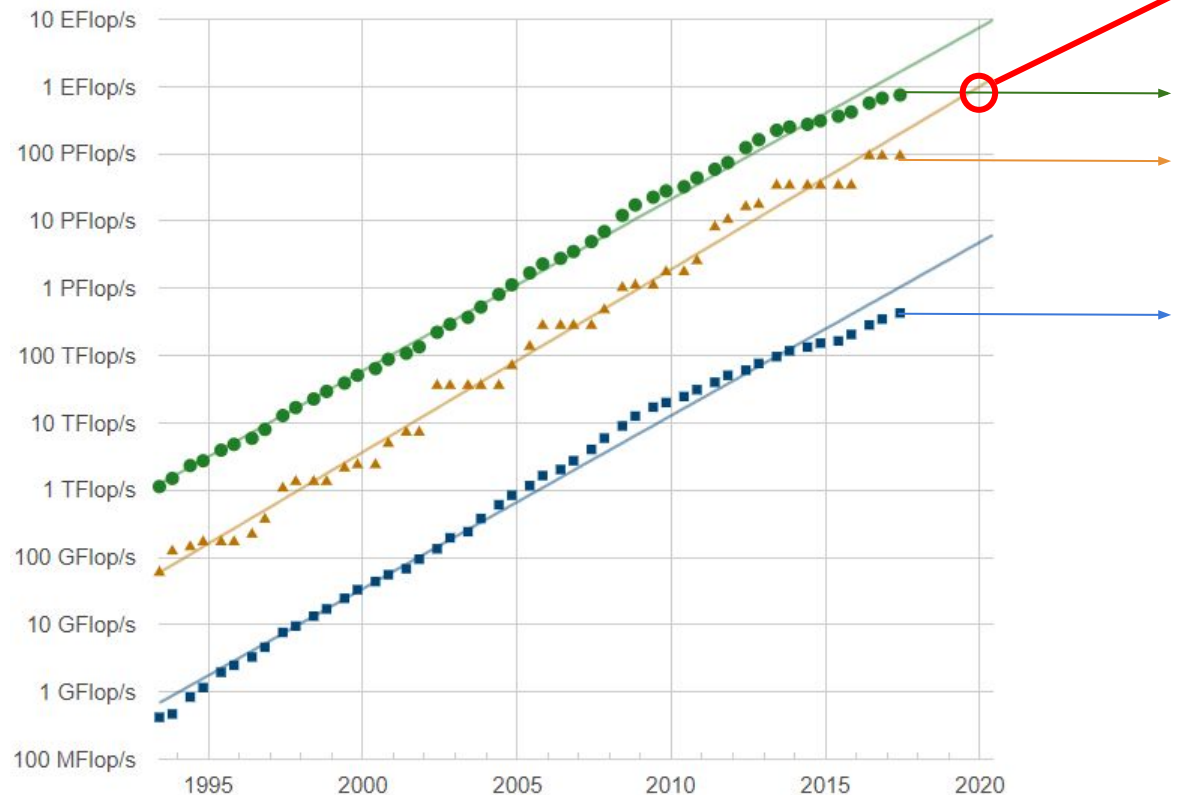
Supercomputadores

- São basicamente clusters de multicores e GPUs conectados por uma rede rápida.



Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Supercomputing Center in Wuxi China	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRPC	10,649,600	93,014.6	125,435.9	15,371
2	National Super Computer Center in Guangzhou China	Tianhe-2 (MilkyWay-2) - TH-IVB- FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
3	Swiss National Supercomputing Centre (CSCS) Switzerland	Piz Daint - Cray XC50, Xeon E5- 2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 Cray Inc.	361,760	19,590.0	25,326.3	2,272
4	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
5	DOE/NNSA/LLNL United States	Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890

Desempenho dos supercomputadores em Flops por ano



O maior supercomputador do mundo atingirá a capacidade de Exaflops em 2020

Soma do poder computacional de todos os computadores do TOP500

Poder computacional do computador rank #1 do TOP500

Poder computacional do computador rank #500 do TOP500