

Ambientes de computação de alto desempenho no LNCC

Roberto Pinto Souto
MCTI/LNCC/CSR - CENAPAD-RJ
rpsouto@lncc.br

24 de Março de 2014

Roteiro

1 Introdução

2 Recursos de *Hardware*

3 Recursos de *Software*

4 Considerações Finais

Roteiro

1 Introdução

2 Recursos de *Hardware*

3 Recursos de *Software*

4 Considerações Finais

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo factível, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de núcleos computacionais ("cores"), distribuídos em processadores convencionais (CPUs) e aceleradores (GPUs). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda sequencialmente, ter o conhecimento das diferentes opções de ferramentas computacionais que dão suporte à uma execução paralela. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de núcleos computacionais ("cores"), distribuídos em processadores convencionais (CPUs) e aceleradores (GPUs). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda sequencialmente, ter o conhecimento das diferentes opções de ferramentas computacionais que dão suporte à uma execução paralela. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de *núcleos computacionais ("cores")*, distribuídos em processadores convencionais (CPUs) e aceleradores (GPUs). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda sequencialmente, ter o conhecimento das diferentes opções de ferramentas computacionais que dão suporte à uma execução paralela. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de *núcleos computacionais ("cores")*, distribuídos em processadores convencionais (*CPUs*) e aceleradores (*GPUs*). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda sequencialmente, ter o conhecimento das diferentes opções de ferramentas computacionais que dão suporte à uma execução paralela. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de *núcleos computacionais ("cores")*, distribuídos em processadores convencionais (*CPUs*) e aceleradores (*GPUs*). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda *sequencialmente*, ter o conhecimento das diferentes opções de ferramentas computacionais que dão suporte à uma execução paralela. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de *núcleos computacionais ("cores")*, distribuídos em processadores convencionais (*CPUs*) e aceleradores (*GPUs*). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda *sequencialmente*, ter o conhecimento das diferentes opções de *ferramentas computacionais* que dão suporte à uma execução paralela. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de *núcleos computacionais ("cores")*, distribuídos em processadores convencionais (*CPUs*) e aceleradores (*GPUs*). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda *sequencialmente*, ter o conhecimento das diferentes opções de *ferramentas computacionais* que dão suporte à uma *execução paralela*. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em computação de alto desempenho disponibilizados no LNCC.

Introdução

Apresentação

Diversas aplicações científicas podem executar um número tão elevado de operações que, a fim de que sejam finalizadas em um tempo *factível*, se faz necessária a utilização de máquinas com cada vez maior poder de processamento. Notadamente, estas são máquinas paralelas que possuem um determinado número de *núcleos computacionais* ("*cores*"), distribuídos em processadores convencionais (*CPUs*) e aceleradores (*GPUs*). É importante, portanto, o usuário que desejar reduzir o tempo de processamento de uma aplicação que roda *sequencialmente*, ter o conhecimento das diferentes opções de *ferramentas computacionais* que dão suporte à uma *execução paralela*. O LNCC dispõe de máquinas paralelas que contém compiladores, bibliotecas e ferramentas que possibilitam ao usuário obter desempenho paralelo dos códigos científicos utilizados. Neste seminário será dada uma visão geral destes recursos e de serviços em *computação de alto desempenho* disponibilizados no LNCC.

Computação de Alto Desempenho

Computação de Alto Desempenho

High Performance Computing (HPC)

Computação de Alto Desempenho

High Performance Computing (HPC)

Definition

“High Performance Computing most generally refers to the practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business.”

FONTE: <http://insidehpc.com/hpc-basic-training/what-is-hpc>

Computação de Alto Desempenho

High Performance Computing (HPC)

Definition

“High Performance Computing most generally refers to the practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business.”

FONTE: <http://insidehpc.com/hpc-basic-training/what-is-hpc>

Definição - *tradução livre*

“Computação de Alto Desempenho, de um modo geral, se refere à prática de agregar poder de processamento a fim de proporcionar um desempenho muito maior do que se poderia obter de um desktop ou estação de trabalho, para resolver grandes problemas em ciência, em engenharia ou comerciais.”

FONTE: <http://insidehpc.com/hpc-basic-training/what-is-hpc>

Instituições que demandam poder computacional

Comerciais

- Ramo financeiro;
- Seguradoras;
- Sistemas industriais;

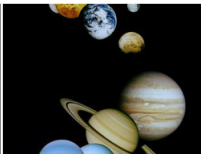
Científicas

- Universidades;
- Centros de pesquisa;
- Indústrias aeronáutica, automobilística, de energia, entre outras;

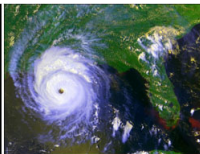
Problemas de Grande Porte



Galaxy Formation



Planetary Movments



Climate Change



Rush Hour Traffic



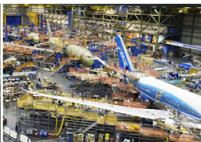
Plate Tectonics



Weather



Auto Assembly



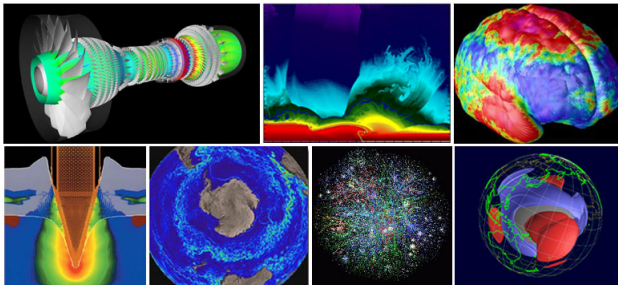
Jet Construction



Drive-thru Lunch

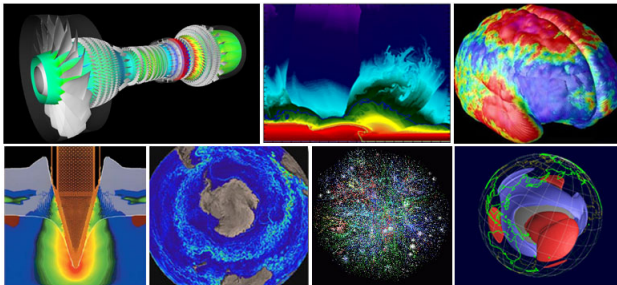
FONTE: https://computing.llnl.gov/tutorials/parallel_comp/

Simulações de Grande Porte



Aplicações Científicas:

Simulações de Grande Porte



Aplicações Científicas:

Atmosfera, ambiental

Física: aplicada, nuclear, da partícula, matéria condensada, fusão, fotônica

Biociência, Biotecnologia, Genética

Química

Geologia

Engenharia Mecânica: indústria aeroespacial, indústria automobilística, energia

Engenharia Elétrica: projeto de circuitos, microeletrônica

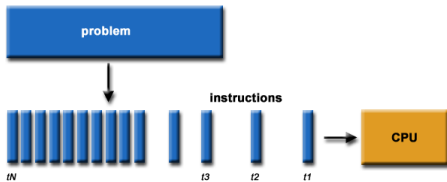
Ciência da Computação, Matemática

Defesa, material bélico

Aplicações Científicas no LNCC

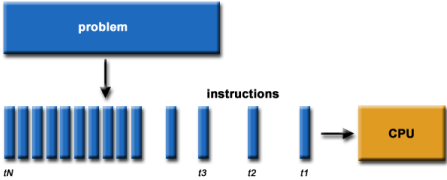
- Previsão do Tempo
- Meta-heurísticas Massivamente Paralelas
- Métodos de Elementos Finitos Multi-escalas
- Predição de Estruturas de Proteínas
- Simulação Paralela de Computação Quântica
- Simulação em Reservatórios de Petróleo
- Bionformática
- Workflow Científico para Modelagem de Biodiversidade
- Gerenciamento e Mineração de Dados em Ciência
- Simulação Computacional do Sistema Cardiovascular Humano

Processamento Paralelo

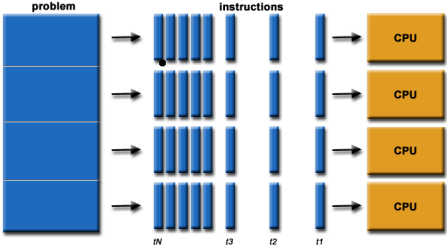


Programa sequencial ou serial

Processamento Paralelo



Programa sequencial ou serial



Programa Paralelo

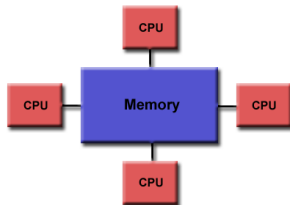
Arquitetura de Computadores Paralelos

Memória compartilhada

Memória Distribuída

Arquitetura de Computadores Paralelos

Memória compartilhada

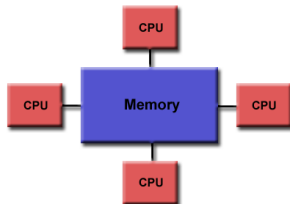


Uniform Memory Access (UMA)

Memória Distribuída

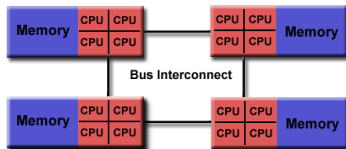
Arquitetura de Computadores Paralelos

Memória compartilhada



Memória Distribuída

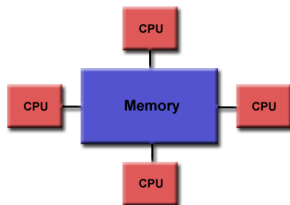
Uniform Memory Access (UMA)



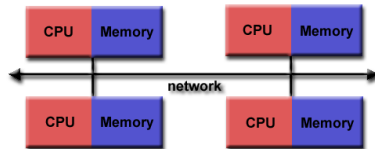
Non-Uniform Memory Access (NUMA)

Arquitetura de Computadores Paralelos

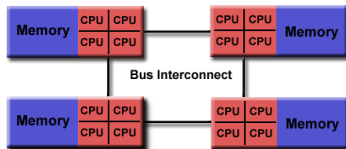
Memória compartilhada



Memória Distribuída



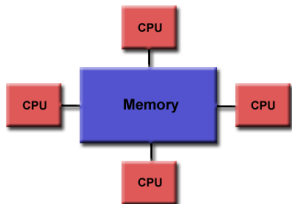
Uniform Memory Access (UMA)



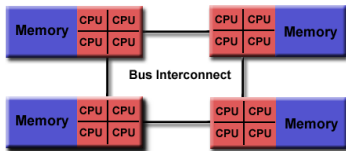
Non-Uniform Memory Access (NUMA)

Arquitetura de Computadores Paralelos

Memória compartilhada

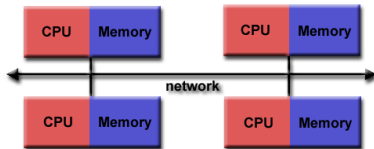


Uniform Memory Access (UMA)

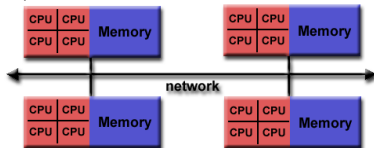


Non-Uniform Memory Access (NUMA)

Memória Distribuída

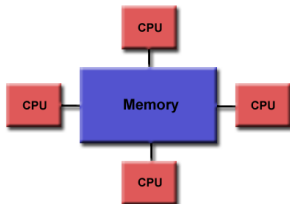


Arquitetura Híbrida

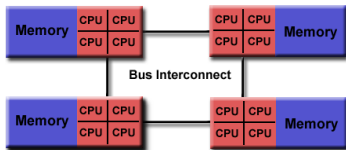


Arquitetura de Computadores Paralelos

Memória compartilhada

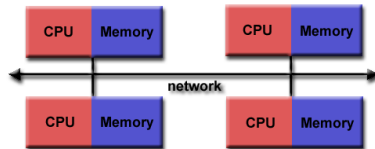


Uniform Memory Access (UMA)

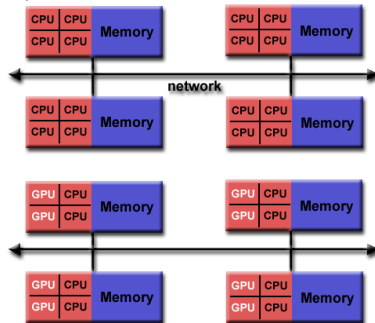


Non-Uniform Memory Access (NUMA)

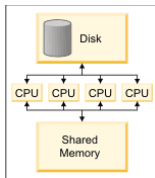
Memória Distribuída



Arquitetura Híbrida

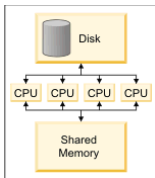


Arquitetura de Computadores Paralelos

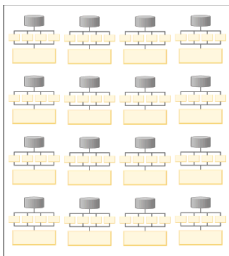


Symmetric Multi-Processing (SMP)

Arquitetura de Computadores Paralelos

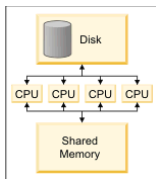


Symmetric Multi-Processing (SMP)

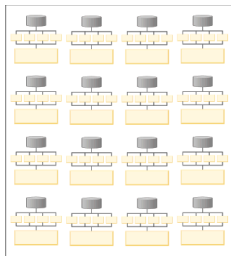


Massively Parallel Processing (MPP)

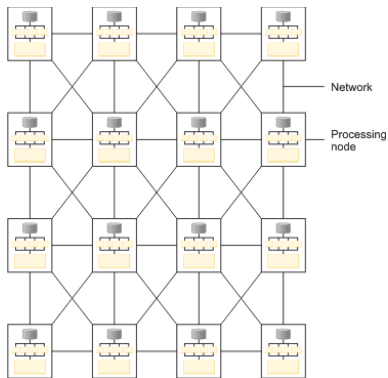
Arquitetura de Computadores Paralelos



Symmetric Multi-Processing (SMP)



Massively Parallel Processing (MPP)



Cluster

Máquinas paralelas

Cluster

Máquinas paralelas

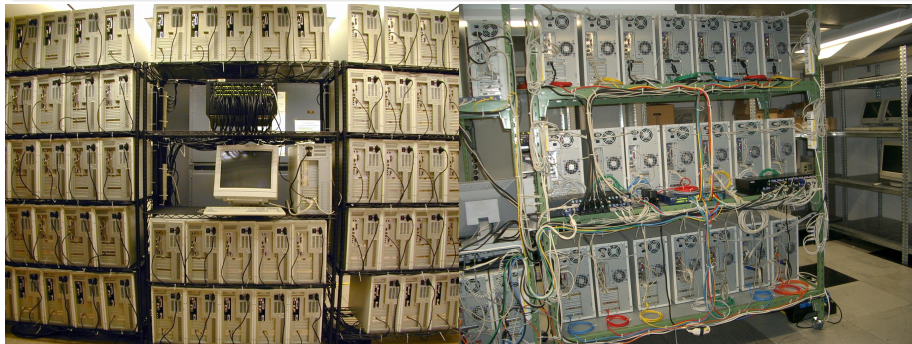
Cluster

- computadores agregados por uma rede de interconexão;

Máquinas paralelas

Cluster

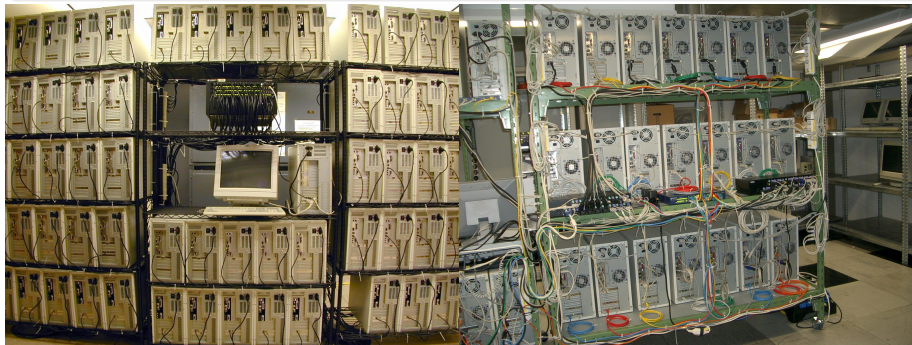
- computadores agregados por uma rede de interconexão;



Máquinas paralelas

Cluster

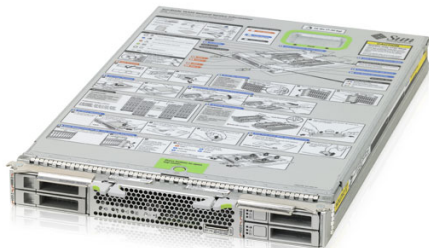
- computadores agregados por uma rede de interconexão;
- máquina de memória distribuída;



Máquinas paralelas

Cluster

- computadores agregados por uma rede de interconexão;
- máquina de memória distribuída;



Cluster

Propriedades

- Boa relação Custo/Desempenho

Cluster

Propriedades

- Boa relação Custo/Desempenho
- Configuração ajustável

Cluster

Propriedades

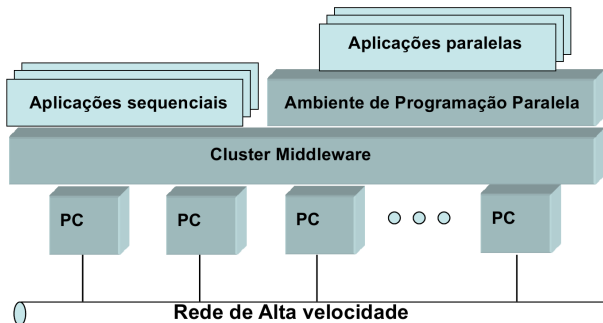
- Boa relação Custo/Desempenho
- Configuração ajustável
- Escalável

Cluster

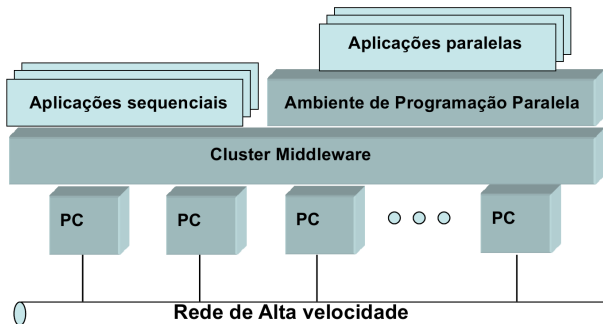
Propriedades

- Boa relação Custo/Desempenho
- Configuração ajustável
- Escalável
- Rápida incorporação de novas tecnologias

Arquitetura de um Cluster

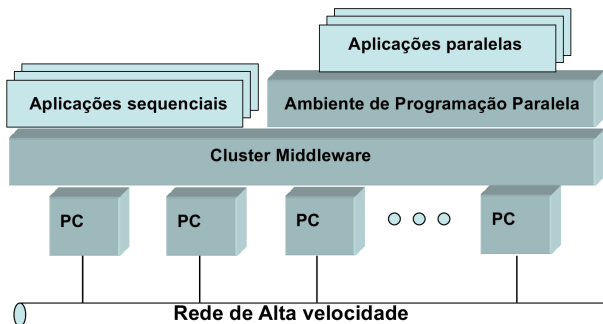


Arquitetura de um Cluster



Rede de Alta Velocidade / Rede de Interconexão

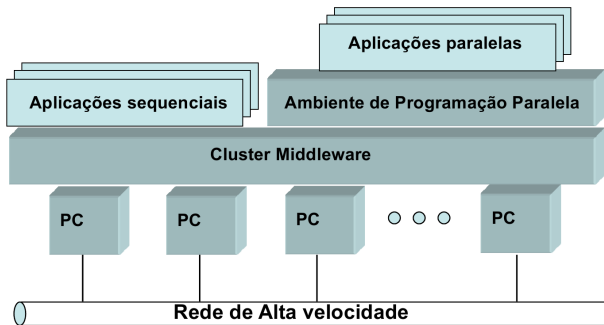
Arquitetura de um Cluster



Rede de Alta Velocidade / Rede de Interconexão

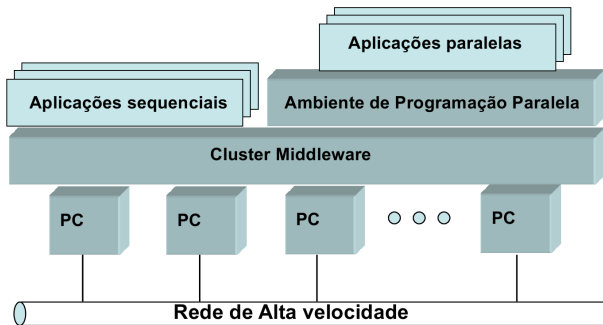
- **Padrões:** Gigabit Ethernet (até 100 Gb/s), Infiniband (até 300 Gb/s) e Proprietárias

Arquitetura de um Cluster



PC / NÓS de Processamento

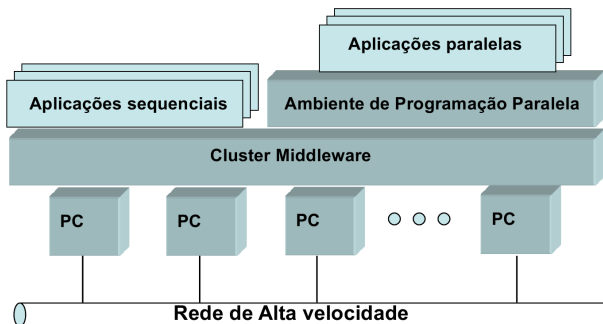
Arquitetura de um Cluster



PC / NÓS de Processamento

- **Arquiteturas:** multi-core (CPUs) e many-core (coprocessadores)

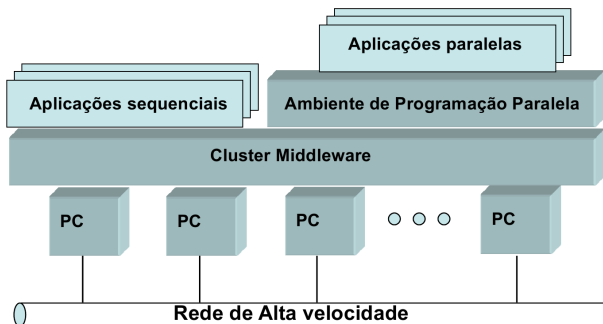
Arquitetura de um Cluster



PC / NÓS de Processamento

- **Arquiteturas:** multi-core (CPUs) e many-core (coprocessadores)
- **multi-core:** Intel Xeon, AMD Opteron, IBM Power

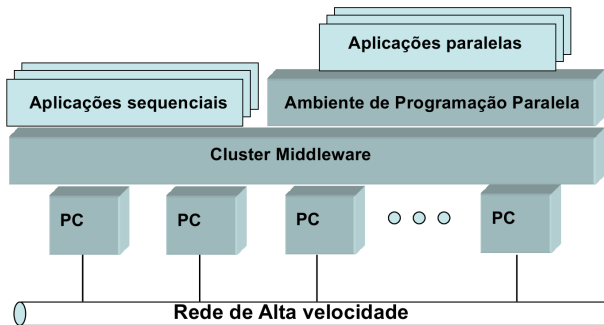
Arquitetura de um Cluster



PC / NÓS de Processamento

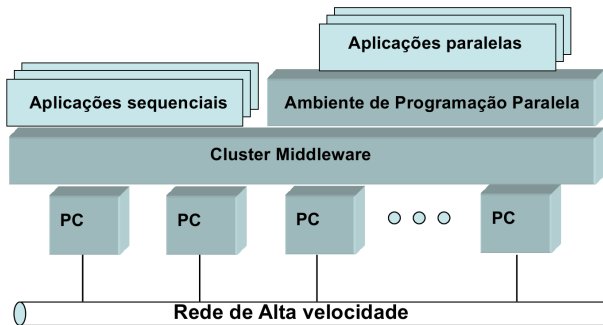
- **Arquiteturas:** multi-core (CPUs) e many-core (coprocessadores)
- **multi-core:** Intel Xeon, AMD Opteron, IBM Power
- **many-core:** Intel Xeon Phi, AMD ATI Radeon, Nvidia Tesla

Arquitetura de um Cluster



Cluster Middleware

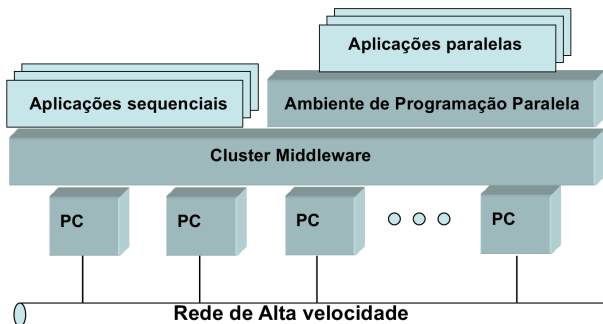
Arquitetura de um Cluster



Cluster Middleware

- **Sistema Operacional:** CentOS, Suse, Debian, etc ...

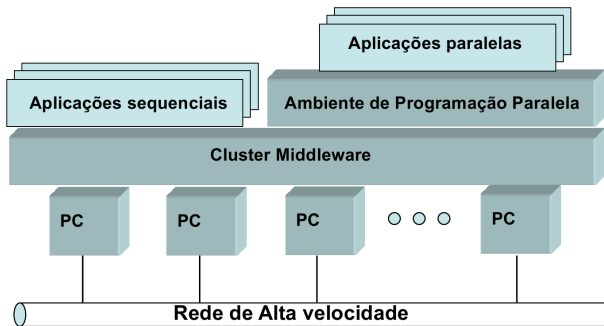
Arquitetura de um Cluster



Cluster Middleware

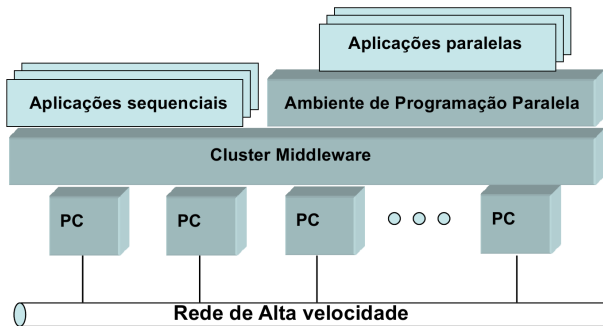
- **Sistema Operacional:** CentOS, Suse, Debian, etc ...
- **Sistema de Gerenciamento de Recursos:** PBS/Torque, SGE, etc...

Arquitetura de um Cluster



Ambiente de Programação Paralela

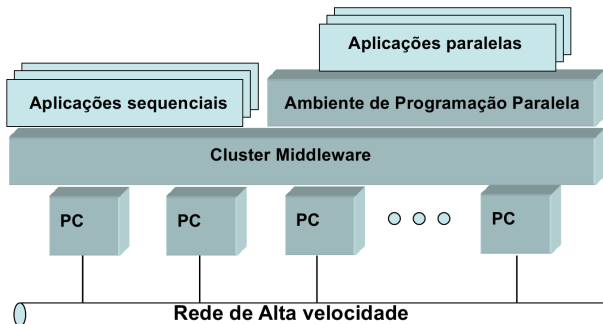
Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Entre os nós / memória distribuída:** MPI (Message Passing Interface)

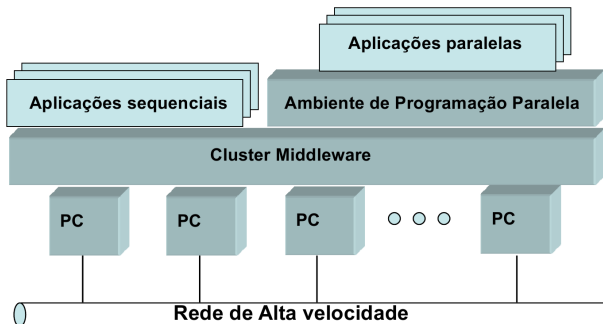
Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Entre os nós / memória distribuída:** MPI (Message Passing Interface)
- **Dentro do nó / memória compartilhada:** pthreads, OpenMP, OpenACC, OpenCL

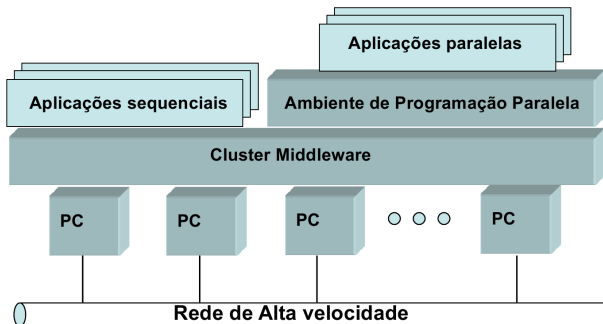
Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Entre os nós / memória distribuída:** MPI (Message Passing Interface) OpenSHMEM*, UPC (Unified Parallel C)*, CAF (Co-Array Fortran)*
- **Dentro do nó / memória compartilhada:** pthreads, OpenMP, OpenACC, OpenCL

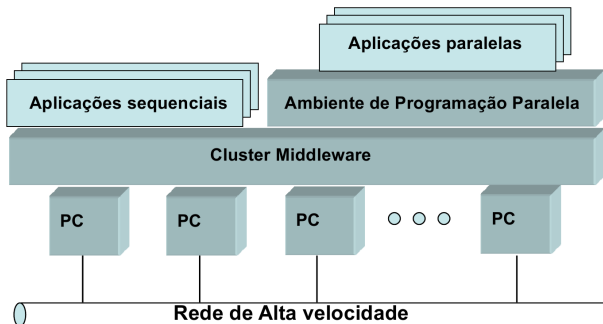
Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Entre os nós / memória distribuída:** MPI (Message Passing Interface) OpenSHMEM*, UPC (Unified Parallel C)*, CAF (Co-Array Fortran)*
- **Dentro do nó / memória compartilhada:** pthreads, OpenMP, OpenACC, OpenCL, CilkPlus*, TBB*, CUDA*

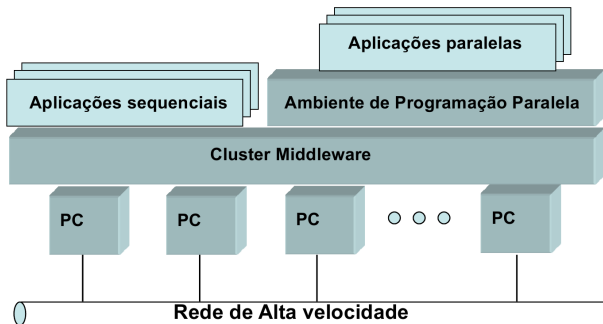
Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Entre os nós / memória distribuída:** MPI (Message Passing Interface)
***PGAS** (Partioned Global Address Space)
- **Dentro do nó / memória compartilhada:**
pthreads, OpenMP, OpenACC, OpenCL, ***“Proprietárias”**

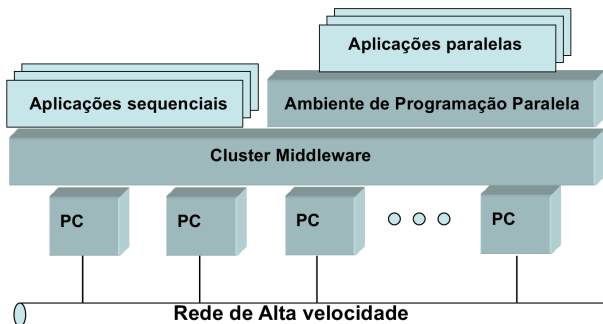
Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Compiladores (*produção*):** GNU, Intel, PGI, Cray

Arquitetura de um Cluster



Ambiente de Programação Paralela

- **Compiladores (*produção*):** GNU, Intel, PGI, Cray
- **Compiladores (*pesquisa*):** Open64/OpenUH
Mercurium/Nanos++ (*BSC*), Rose (*LLNL*)

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;
- **Disponibilidade:** gerenciamento de falhas;

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;
- **Disponibilidade:** gerenciamento de falhas;
- **Imagem única do sistema:** parece ao usuário como se fosse um único sistema;

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;
- **Disponibilidade:** gerenciamento de falhas;
- **Imagem única do sistema:** parece ao usuário como se fosse um único sistema;
- **Comunicação rápida:** redes e protocolos de comunicação;

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;
- **Disponibilidade:** gerenciamento de falhas;
- **Imagem única do sistema:** parece ao usuário como se fosse um único sistema;
- **Comunicação rápida:** redes e protocolos de comunicação;
- **Balanceamento de carga:** CPU, rede, memória, discos;

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;
- **Disponibilidade:** gerenciamento de falhas;
- **Imagem única do sistema:** parece ao usuário como se fosse um único sistema;
- **Comunicação rápida:** redes e protocolos de comunicação;
- **Balanceamento de carga:** CPU, rede, memória, discos;
- **Gerenciamento:** administração e controle;

Cluster

Desafios

- **Escalabilidade:** física e de aplicação;
- **Disponibilidade:** gerenciamento de falhas;
- **Imagem única do sistema:** parece ao usuário como se fosse um único sistema;
- **Comunicação rápida:** redes e protocolos de comunicação;
- **Balanceamento de carga:** CPU, rede, memória, discos;
- **Gerenciamento:** administração e controle;
- **Aplicabilidade:** aplicações voltadas para o cluster.

Cluster

Medidas de Desempenho

- **Vazão (throughput):** taxa na qual os pedidos são atendidos pelo sistema

Cluster

Medidas de Desempenho

- **Vazão (throughput):** taxa na qual os pedidos são atendidos pelo sistema
- **Utilização:** fração do tempo em que o recurso permanece ocupado

Cluster

Medidas de Desempenho

- **Vazão (throughput):** taxa na qual os pedidos são atendidos pelo sistema
- **Utilização:** fração do tempo em que o recurso permanece ocupado
- **Tempo de resposta (latência):** tempo decorrido entre o pedido e o início/conclusão da realização do serviço

Cluster

Medidas de Desempenho

- **Vazão (throughput):** taxa na qual os pedidos são atendidos pelo sistema
- **Utilização:** fração do tempo em que o recurso permanece ocupado
- **Tempo de resposta (latência):** tempo decorrido entre o pedido e o início/conclusão da realização do serviço
- **Escalabilidade:** um sistema é dito escalável quando a eficiência se mantém próxima a ideal com o aumento do número p de processadores aplicado à solução do problema cresce

Cluster

Medidas de Desempenho

- T_s : tempo de execução do programa sequencial (um processador)

Cluster

Medidas de Desempenho

- T_s : tempo de execução do programa sequencial (um processador)
- T_p : tempo de execução do programa paralelo em p de processadores

Cluster

Medidas de Desempenho

- T_s : tempo de execução do programa sequencial (um processador)
- T_p : tempo de execução do programa paralelo em p de processadores
- **Speed-up ou ganho (S_p)**: razão entre T_s e T_p

$$S_p = \frac{T_s}{T_p}$$

Cluster

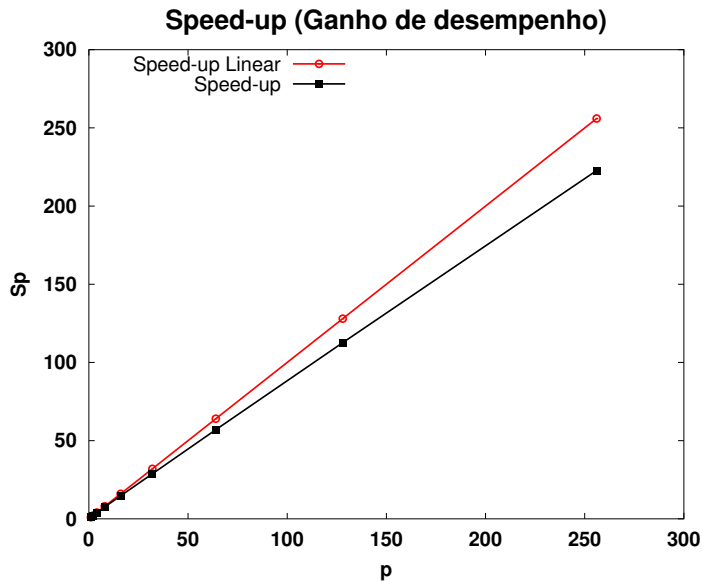
Medidas de Desempenho

- T_s : tempo de execução do programa sequencial (um processador)
- T_p : tempo de execução do programa paralelo em p de processadores
- **Speed-up ou ganho (S_p)**: razão entre T_s e T_p

$$S_p = \frac{T_s}{T_p}$$

- **Speed-up ideal ou linear:** $S_p = p$

Cluster



Cluster

Medidas de Desempenho

- **Eficiência** E_p : razão entre o **speed-up** obtido, pelo número p de processadores

$$E_p = \frac{S_p}{p}$$

Cluster

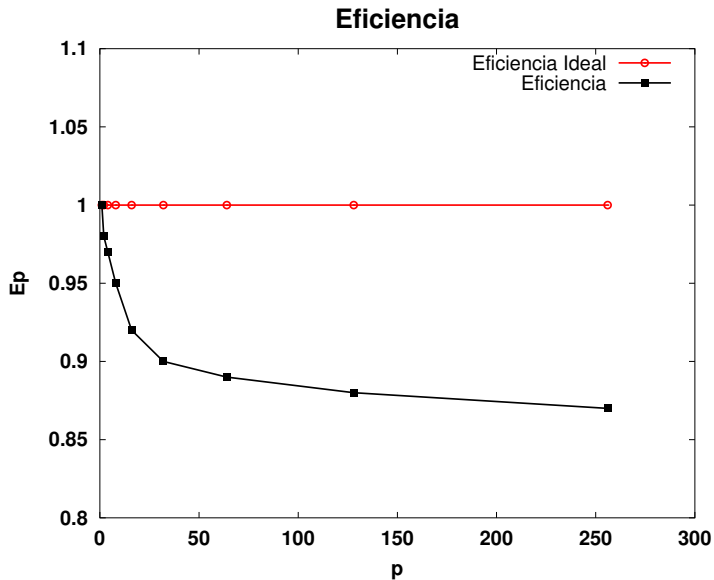
Medidas de Desempenho

- **Eficiência** E_p : razão entre o **speed-up** obtido, pelo número p de processadores

$$E_p = \frac{S_p}{p}$$

- **Eficiência ideal:** $E_p = 1$

Cluster



Cluster

Métricas de Desempenho Paralelo

- **GFLOP/S:** bilhões (Giga) de operações de ponto flutuante por segundo

Cluster

Métricas de Desempenho Paralelo

- **GFLOP/S:** bilhões (Giga) de operações de ponto flutuante por segundo
- **Pico teórico de desempenho:** ocorre no caso em que a máquina não fizesse nada além de operações numéricas

Cluster

Métricas de Desempenho Paralelo

- **GFLOP/S:** bilhões (Giga) de operações de ponto flutuante por segundo
- **Pico teórico de desempenho:** ocorre no caso em que a máquina não fizesse nada além de operações numéricas
- **Benchmarks:** programas desenvolvidos para determinar as métricas de desempenho da máquinas. Exemplos: LINPACK, NASA-NPB

Cluster

LINKPACK Benchmark

- Surgiu como pacote de solução de sistemas lineares ($Ax = b$)

Cluster

LINKPACK Benchmark

- Surgiu como pacote de solução de sistemas lineares ($Ax = b$)
- Foi utilizado como *benchmark* inicialmente para predição de tempo de execução;

Cluster

LINKPACK Benchmark

- Surgiu como pacote de solução de sistemas lineares ($Ax = b$)
- Foi utilizado como *benchmark* inicialmente para predição de tempo de execução;
- Relacionou-se o tempo de execução de 23 computadores para um problema de ordem 100 (LINKPACK100)

Cluster

LINKPACK Benchmark

- Surgiu como pacote de solução de sistemas lineares ($Ax = b$)
- Foi utilizado como *benchmark* inicialmente para predição de tempo de execução;
- Relacionou-se o tempo de execução de 23 computadores para um problema de ordem 100 (LINKPACK100)
- Esta é a origem ranking mundial de supercomputadores **TOP500**
<http://www.top500.org/>

Cluster

LINKPACK Benchmark

- Surgiu como pacote de solução de sistemas lineares ($Ax = b$)
- Foi utilizado como *benchmark* inicialmente para predição de tempo de execução;
- Relacionou-se o tempo de execução de 23 computadores para um problema de ordem 100 (LINKPACK100)
- Esta é a origem ranking mundial de supercomputadores **TOP500**
<http://www.top500.org/>
- O TOP500 atualmente conta com dados de mais de 1300 sistemas de computação.

Cluster

LINKPACK Benchmark

- Surgiu como pacote de solução de sistemas lineares ($Ax = b$)
- Foi utilizado como *benchmark* inicialmente para predição de tempo de execução;
- Relacionou-se o tempo de execução de 23 computadores para um problema de ordem 100 (LINKPACK100)
- Esta é a origem ranking mundial de supercomputadores **TOP500**
<http://www.top500.org/>
- O TOP500 atualmente conta com dados de mais de 1300 sistemas de computação.
- High Performance LINKPACK $n \times n$ (HPL):
<http://www.netlib.org/benchmark/hpl>

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;
- Esforço humano em otimizar o código;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;
- Esforço humano em otimizar o código;
- Sistema Operacional;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;
- Esforço humano em otimizar o código;
- Sistema Operacional;
- Hardware;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;
- Esforço humano em otimizar o código;
- Sistema Operacional;
- Hardware;
- Idade do compilador;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;
- Esforço humano em otimizar o código;
- Sistema Operacional;
- Hardware;
- Idade do compilador;
- Habilidade do compilador em gerar código otimizado;

LINKPACK Benchmark

Fatores que afetam o desempenho

- Tipo de aplicação;
- Algoritmo;
- Tamanho do problema;
- Nível da linguagem de programação;
- Implementação;
- Esforço humano em otimizar o código;
- Sistema Operacional;
- Hardware;
- Idade do compilador;
- Habilidade do compilador em gerar código otimizado;
- Qualidade do núcleo de cálculo (BLAS otimizada);

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):
<http://www.netlib.org/blas>

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):

<http://www.netlib.org/blas>

Nível 1: operações vetor-vetor

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):

<http://www.netlib.org/blas>

Nível 1: operações vetor-vetor

Nível 2: operações matriz-vetor

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):

<http://www.netlib.org/blas>

Nível 1: operações vetor-vetor

Nível 2: operações matriz-vetor

Nível 3: operações matriz-matriz

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):

<http://www.netlib.org/blas>

Nível 1: operações vetor-vetor

Nível 2: operações matriz-vetor

Nível 3: operações matriz-matriz

- Procurar utilizar BLAS otimizadas:

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):
<http://www.netlib.org/blas>
Nível 1: operações vetor-vetor
Nível 2: operações matriz-vetor
Nível 3: operações matriz-matriz
- Procurar utilizar BLAS otimizadas:
ATLAS: gerador gratuito de BLAS otimizada para qualquer sistema
<http://www.netlib.org/atlas>

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):
<http://www.netlib.org/blas>
Nível 1: operações vetor-vetor
Nível 2: operações matriz-vetor
Nível 3: operações matriz-matriz
- Procurar utilizar BLAS otimizadas:
ATLAS: gerador gratuito de BLAS otimizada para qualquer sistema
<http://www.netlib.org/atlas>
Intel MKL (Intel Math Kernel Library)

Cluster

Núcleo de cálculo do LINPACK

- BLAS (Basic Linear Algebra Subprograms):
<http://www.netlib.org/blas>
Nível 1: operações vetor-vetor
Nível 2: operações matriz-vetor
Nível 3: operações matriz-matriz
- Procurar utilizar BLAS otimizadas:
ATLAS: gerador gratuito de BLAS otimizada para qualquer sistema
<http://www.netlib.org/atlas>
Intel MKL (Intel Math Kernel Library)
Goto BLAS

Cluster

Em que consiste a lista TOP500

- R_{max} : Desempenho em Gflop/s para o maior problema executado;

Cluster

Em que consiste a lista TOP500

- R_{max} : Desempenho em Gflop/s para o maior problema executado;
- N_{max} : A dimensão do maior problema executado

Cluster

Em que consiste a lista TOP500

- R_{max} : Desempenho em Gflop/s para o maior problema executado;
- N_{max} : A dimensão do maior problema executado
- $N_{1/2}$: O tamanho do problema onde foi alcançado metade de R_{max} ;

Cluster

Em que consiste a lista TOP500

- R_{max} : Desempenho em Gflop/s para o maior problema executado;
- N_{max} : A dimensão do maior problema executado
- $N_{1/2}$: O tamanho do problema onde foi alcançado metade de R_{max} ;
- R_{peak} : O pico teórico de desempenho da máquina

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);
 $FLOPS \approx \#cores \times freq \times flop/cycle$

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);

$$FLOPS \approx \#cores \times freq \times flop/cycle$$

$$FLOPS \approx 8 \times 2.7GHz \times 8$$

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);

$$FLOPS \approx \#cores \times freq \times flop/cycle$$

$$FLOPS \approx 8 \times 2.7GHz \times 8$$

$$FLOPS \approx \mathbf{172.8Gflop/s}$$

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);
 $FLOPS \approx \#cores \times freq \times flop/cycle$
 $FLOPS \approx 8 \times 2.7GHz \times 8$
 $FLOPS \approx \mathbf{172.8Gflop/s}$
- Se integarmos **18 432** nós com esta configuração

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);
 $FLOPS \approx \#cores \times freq \times flop/cycle$
 $FLOPS \approx 8 \times 2.7GHz \times 8$
 $FLOPS \approx \mathbf{172.8Gflop/s}$
- Se integrarmos **18 432** nós com esta configuração
- Temos então um cluster com **3 185 049.6 GFlop/s**

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);
 $FLOPS \approx \#cores \times freq \times flop/cycle$
 $FLOPS \approx 8 \times 2.7GHz \times 8$
 $FLOPS \approx \mathbf{172.8Gflop/s}$
- Se integrarmos **18 432** nós com esta configuração
- Temos então um cluster com **3 185 049.6 GFlop/s**
- ou **3 185 TFlop/s**

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);
 $FLOPS \approx \#cores \times freq \times flop/cycle$
 $FLOPS \approx 8 \times 2.7GHz \times 8$
 $FLOPS \approx 172.8Gflop/s$
- Se integrarmos **18 432** nós com esta configuração
- Temos então um cluster com **3 185 049.6 GFlop/s**
- ou **3 185 TFlop/s**
- ou **3.2 PFlop/s** de pico teórico

Desempenho máximo teórico R_{peak}

Intel Xeon E5-2680 (Sandy Bridge):

- Possui 8 *cores* ($\#cores = 8$)
- Frequência de clock de 2.7GHz ($freq = 2.7GHz$);
- Executando **8** operações de ponto flutuante de **precisão dupla** por ciclo de clock ($\#flop/cycle = 8$);
 $FLOPS \approx \#cores \times freq \times flop/cycle$
 $FLOPS \approx 8 \times 2.7GHz \times 8$
 $FLOPS \approx 172.8Gflop/s$
- Se integrarmos **18 432** nós com esta configuração
- Temos então um cluster com **3 185 049.6 GFlop/s**
- ou **3 185 TFlop/s**
- ou **3.2 PFlop/s** de pico teórico
- Esta é a configuração do supercomputador número 10 no TOP500:
SuperMUC (Alemanha)

TOP500: #10 SuperMUC (Alemanha)

SuperMUC - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR



Configuração

- **147 456 cores** distribuídos em nós contendo CPU Intel Xeon E5-2680 8-Core
- $R_{max} = 2.9$ PFlop/s
- $R_{peak} = 3.2$ PFlop/s

Site:	Leibniz Rechenzentrum
System URL:	http://www.lrz.de/services/compute/supermuc/
Manufacturer:	IBM
Cores:	147,456
Linpack Performance (Rmax)	2,897.0 TFlop/s
Theoretical Peak (Rpeak)	3,185.1 TFlop/s
Power:	3,422.67 kW
Memory:	
Interconnect:	Infiniband FDR
Operating System:	Linux

TOP500: #2 TITAN (EUA)

Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x



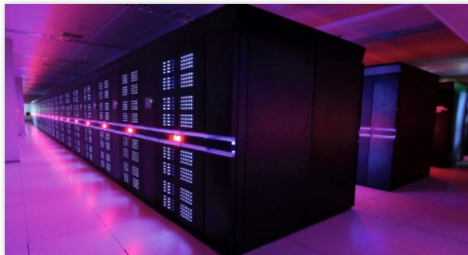
Site:	DOE/SC/Oak Ridge National Laboratory
System URL:	http://www.olcf.ornl.gov/titan/
Manufacturer:	Cray Inc.
Cores:	560,640
Linpack Performance (Rmax)	17,590.0 TFlop/s
Theoretical Peak (Rpeak)	27,112.5 TFlop/s
Power:	8,209.00 kW
Memory:	710,144 GB
Interconnect:	Cray Gemini interconnect
Operating System:	Cray Linux Environment

Configuração

- **147 456 cores** distribuídos em nós contendo CPU AMD Opteron 6274 16-Core
- e em GPU Nvidia K20
- $R_{max} = 17.6$ PFlop/s
- $R_{peak} = 27.1$ PFlop/s

TOP500: #1 Tianhe-2 (China)

Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C
2.200GHz, TH Express-2, Intel Xeon Phi 31S1P



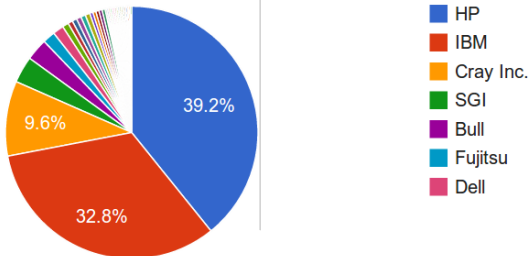
Site:	National Super Computer Center in Guangzhou
Manufacturer:	NUDT
Cores:	3,120,000
Linpack Performance (Rmax)	33,862.7 TFlop/s
Theoretical Peak (Rpeak)	54,902.4 TFlop/s
Power:	17,808.00 kW
Memory:	1,024,000 GB
Interconnect:	TH Express-2
Operating System:	Kylin Linux
Compiler:	icc
Math Library:	Intel MKL-11.0.0
MPI:	MPICH2 with a customized GLEX channel

Configuração

- **3 120 000 cores** distribuídos em nós contendo CPU Intel Xeon E5-2692 12-Core
- e em coprocessador Intel Xeon Phi 31S1P
- $R_{max} = 33.9$ PFlop/s
- $R_{peak} = 54.9$ PFlop/s

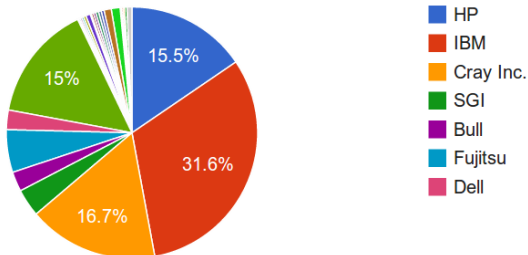
Principais Fabricantes (*Manufacturer/Vendors*):

Vendors System Share

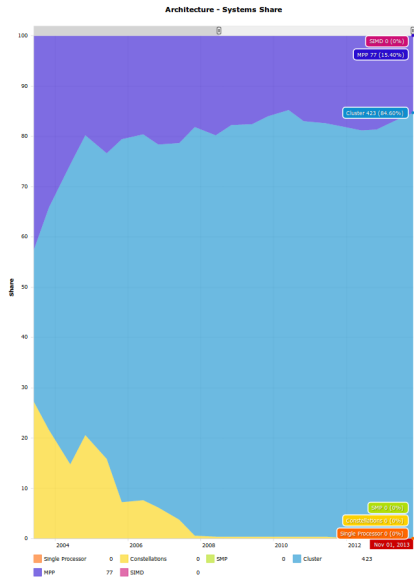


Principais Fabricantes (*Manufacturer/Vendors*):

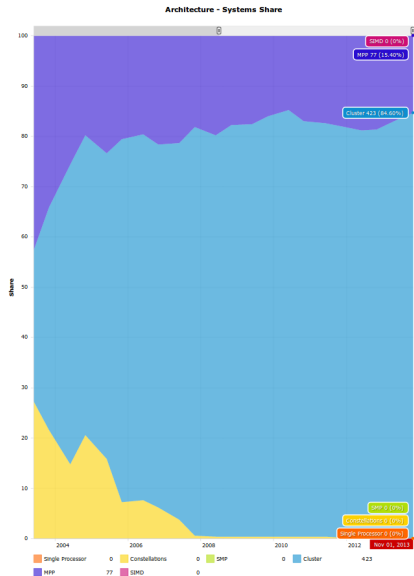
Vendors Performance Share



TOP500: Arquiteturas



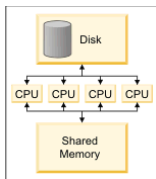
TOP500: Arquiteturas



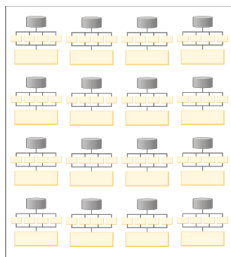
CLUSTERS: 84.60%

MPP : 15.40%

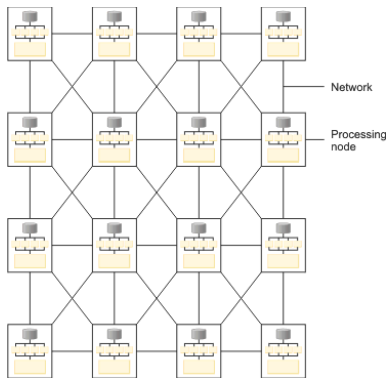
Arquitetura de Computadores Paralelos



Symmetric Multi-Processing (SMP)



Massively Parallel Processing (MPP)



Cluster

Brasil no TOP500

Perspectivas

Perspectivas

- O LNCC está em negociação para receber um cluster da Bull com cerca de 2 PFlop/s de R_{peak}

Perspectivas

- O LNCC está em negociação para receber um cluster da Bull com cerca de 2 PFlop/s de R_{peak}
- Potencialmente, este sistema estaria hoje entre os 20 primeiros no TOP500 (R_{max})

Perspectivas

- O LNCC está em negociação para receber um cluster da Bull com cerca de 2 PFlop/s de R_{peak}
- Potencialmente, este sistema estaria hoje entre os 20 primeiros no TOP500 (R_{max})
- Centro de Pesquisa em HPC da Bull a ser instalado em Petrópolis-RJ

Roteiro

1 Introdução

2 Recursos de *Hardware*

3 Recursos de *Software*

4 Considerações Finais

Clusters no LNCC

Nome	#cores	#cores	R_{peak}	RAM	HD
	CPU(nós)	GPU(nós)	TFlop/s	TBytes	TBytes
Sun Blade X6250	576(30)	0	6.5	1.20	47
SGI Altix XE 340	288(30)	17152(11)	16.2	0.63	129
SGI Altix ICE 8400	300(25)	0	3.2	1.20	12
SGI Altix XE 1300	1128(94)	10752(12)	12.0	N/D	47
Bull bullx	1392(100)	3584(4)	15	3.00	50

Clusters no LNCC: **disponibilizados** pelo CENAPAD-RJ

Nome	CPU	GPU	R_{peak}	RAM	HD
	#cores(nós)	#cores(nós)	TFlop/s	TBytes	TBytes
Sun Blade X6250	576(72)	0	6.5	1.20	47
SGI Altix XE 340	288(30)	17152(11)	16.2	0.63	129
SGI Altix ICE 8400	300(25)	0	3.2	1.20	12
SGI Altix XE 1300	1128(94)	10752(12)	12.0	N/D	47
Bull bullx	1392(100)	3584(4)	15	3.00	50

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **SunHPC**

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**
 - Sistema Operacional: Centos 5.5

- **SunHPC**

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- Sistema Operacional: Centos 5.5

- **SunHPC**

- Sistema Operacional: Centos 5.5

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2

- **SunHPC**

- Sistema Operacional: Centos 5.5

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2

- **SunHPC**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

● **Altix-XE**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;

● **SunHPC**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

● **Altix-XE**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;

● **SunHPC**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

● **Altix-XE**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;
- <http://www.lncc.br/altix-xe>

● **SunHPC**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

● **Altix-XE**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;
- <http://www.lncc.br/altix-xe>

● **SunHPC**

- Sistema Operacional: Centos 5.5
- Gerenciador de Recursos: SGE 6.2
- Rede de Interconexão: Infiniband;
- <http://www.lncc.br/sunhpc>

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **SunHPC**

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**
 - é formado por **30** nós;

- **SunHPC**

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- é formado por **30** nós;

- **SunHPC**

- é formado por **72** nós;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- é formado por **30** nós;
- os nós possuem **até 16** núcleos computacionais;

- **SunHPC**

- é formado por **72** nós;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- é formado por **30** nós;
- os nós possuem **até 16** núcleos computacionais;

- **SunHPC**

- é formado por **72** nós;
- os nós possuem **2 Intel Xeon E5440 2.83GHz Quad Core**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- é formado por **30** nós;
- os nós possuem **até 16** núcleos computacionais;
- **11** desses nós possuem também **GPUs**;

- **SunHPC**

- é formado por **72** nós;
- os nós possuem **2 Intel Xeon E5440 2.83GHz Quad Core**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- é formado por **30** nós;
- os nós possuem **até 16** núcleos computacionais;
- **11** desses nós possuem também **GPUs**;

- **SunHPC**

- é formado por **72** nós;
- os nós possuem **2 Intel Xeon E5440 2.83GHz Quad Core**;
- **nenhum** nó possui **GPU**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- é formado por **30** nós;
- os nós possuem **até 16** núcleos computacionais;
- **11** desses nós possuem também **GPUs**;
- $R_{peak} = 16.2$ TFlop/s

- **SunHPC**

- é formado por **72** nós;
- os nós possuem **2 Intel Xeon E5440 2.83GHz Quad Core**;
- **nenhum** nó possui **GPU**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

● **Altix-XE**

- é formado por **30** nós;
- os nós possuem **até 16** núcleos computacionais;
- **11** desses nós possuem também **GPUs**;
- $R_{peak} = 16.2$ TFlop/s

● **SunHPC**

- é formado por **72** nós;
- os nós possuem **2 Intel Xeon E5440 2.83GHz Quad Core**;
- **nenhum** nó possui **GPU**;
- $R_{peak} = 6.5$ TFlop/s

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**
 - **30** nós;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**
 - **30** nós;
 - **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;
- **4** nós possuem **2 Intel Xeon E5-2660 2.20GHz Octo-Core = 64 cores**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;
- **4** nós possuem **2 Intel Xeon E5-2660 2.20GHz Octo-Core = 64 cores**;
- **CPU cores: 288**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;
- **4** nós possuem **2 Intel Xeon E5-2660 2.20GHz Octo-Core = 64 cores**;
- **CPU cores: 288**;
- **6** nós \times **2 Tesla C2050** \times **448 cores = 5376 cores**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;
- **4** nós possuem **2 Intel Xeon E5-2660 2.20GHz Octo-Core = 64 cores**;
- **CPU cores: 288**;
- **6** nós × **2 Tesla C2050 × 448 cores = 5376 cores**;
- **1** nó × **4 Tesla C2050 × 448 cores = 1792 cores**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;
- **4** nós possuem **2 Intel Xeon E5-2660 2.20GHz Octo-Core = 64 cores**;
- **CPU cores: 288**;
- **6** nós × **2 Tesla C2050 × 448 cores = 5376 cores**;
- **1** nó × **4 Tesla C2050 × 448 cores = 1792 cores**;
- **4** nós × **1 Tesla K20m × 2496 cores = 9984 cores**;

Clusters no LNCC: disponibilizados pelo CENAPAD-RJ

- **Altix-XE**

- **30** nós;
- **19** nós possuem **2 Intel Xeon E5520 2.27GHz Quad-Core = 152 cores**;
- **6** nós possuem **2 Intel Xeon X5650 2.66GHz Hexa-Core = 72 cores**;
- **4** nós possuem **2 Intel Xeon E5-2660 2.20GHz Octo-Core = 64 cores**;
- **CPU cores: 288**;
- **6** nós × **2 Tesla C2050 × 448 cores = 5376 cores**;
- **1** nó × **4 Tesla C2050 × 448 cores = 1792 cores**;
- **4** nós × **1 Tesla K20m × 2496 cores = 9984 cores**;
- **GPU cores: 17152**

Arquitetura das CPUs

Arquitetura das CPUs (cont.)

Sun Blade X6250: 2 Intel Xeon E5440 Quad-Core

Saida da aplicacao de analise de desempenho LIKWID: \$ `likwid-topology -g`

Socket 0:

0	2	4	6
32kB	32kB	32kB	32kB
6MB	6MB		

Socket 1:

1	3	5	7
32kB	32kB	32kB	32kB
6MB	6MB		

Arquitetura das CPUs

Arquitetura das CPUs

Arquitetura das CPUs (cont.)

SGI Altix XE 340 (*node41*): 2 Intel Xeon E5-2660 Octo-Core

Socket 0:

0	16	1	17	2	18	3	19	4	20	5	21	6	22	7	23
32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB
256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB
20MB															

Socket 1:

8	24	9	25	10	26	11	27	12	28	13	29	14	30	15	31
32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB	32kB
256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB	256kB
20MB															

Roteiro

1 Introdução

2 Recursos de *Hardware*

3 Recursos de *Software*

4 Considerações Finais

Recursos de Software

Nível de experiência do usuário em programação paralela

Recursos de Software

Nível de experiência do usuário em programação paralela

Usar aplicação paralela

Recursos de Software

Nível de experiência do usuário em programação paralela

Usar aplicação paralela

não necessita de conhecimento algum de programação;

Recursos de Software

Nível de experiência do usuário em programação paralela

Usar aplicação paralela

não necessita de conhecimento algum de programação;

Usar biblioteca paralela

Recursos de Software

Nível de experiência do usuário em programação paralela

Usar aplicação paralela

não necessita de conhecimento algum de programação;

Usar biblioteca paralela

necessita de conhecimentos básicos de programação;

Recursos de Software

Nível de experiência do usuário em programação paralela

Usar aplicação paralela

não necessita de conhecimento algum de programação;

Usar biblioteca paralela

necessita de conhecimentos básicos de programação;

Desenvolver código-fonte paralelo

Recursos de Software

Nível de experiência do usuário em programação paralela

Usar aplicação paralela

não necessita de conhecimento algum de programação;

Usar biblioteca paralela

necessita de conhecimentos básicos de programação;

Desenvolver código-fonte paralelo

necessita de suficiente conhecimento em programação paralela.

Aplicações Científicas

Portais SINAPAD <https://www.lncc.br/sinapad/portais.php>

Os portais científicos do SINAPAD permitem a execução, via interface Web, de aplicações científicas diversas nos clusters dos centros que compõem o sistema nacional. São 10 as aplicações que podem ser executadas via portal:

ACES III (Advanced Concepts in Electronic Structure III): cálculo de estruturas eletrônicas em química quântica;

BRAMS (Brazilian developments on the Regional Atmospheric Modelling System): previsão numérica de tempo em meso-escala (escala regional);

CAM (Community Atmosphere Model): previsão numérica de tempo em escala global (todo planeta);

Aplicações Científicas

Portais SINAPAD <https://www.lncc.br/sinapad/portais.php>

Os portais científicos do SINAPAD permitem a execução, via interface Web, de aplicações científicas diversas nos clusters dos centros que compõem o sistema nacional. São 10 as aplicações que podem ser executadas via portal:

DANCE (Distributed Assessment of Network CEnterality): calcula a medida de centralidade em redes que classifica os nós pela sua proximidade com todos os demais nós na rede;

DockThor: é um servidor gratuito de atracamento molecular receptor-ligante;

GdfidL: executa simulações eletromagnéticas tridimensionais nos domínios do tempo e frequência, através do Método de Diferenças Finitas;

PrimTest: implementa o teste probabilístico de primalidade de Muller-Rabin;

Aplicações Científicas

Portais SINAPAD <https://www.lncc.br/sinapad/portais.php>

Os portais científicos do SINAPAD permitem a execução, via interface Web, de aplicações científicas diversas nos clusters dos centros que compõem o sistema nacional. São 10 as aplicações que podem ser executadas via portal:

ProFrager: é uma ferramenta de geração de bibliotecas de fragmentos de proteínas;

SPiNMe: é um ambiente computacional em que a especificação e prototipagem de métodos baseados em elementos finitos (FEM), volumes finitos (FVM) e diferenças finitas (DFM);

TrueRNG: explora um processo óptico quântico para gerar fluxos de bits verdadeiramente aleatórios.

Aplicações Científicas

Workflow Científico

Além dos portais SINAPAD, outra abordagem que se destaca no LNCC são os **Workflows Científicos**. Um workflow científico consiste da especificação de um conjunto de aplicações científicas a serem executadas e suas dependências mútuas.

- Segue um ciclo de vida análogo ao dos experimentos científicos computacionais:

Aplicações Científicas

Workflow Científico

Além dos portais SINAPAD, outra abordagem que se destaca no LNCC são os **Workflows Científicos**. Um workflow científico consiste da especificação de um conjunto de aplicações científicas a serem executadas e suas dependências mútuas.

- Segue um ciclo de vida análogo ao dos experimentos científicos computacionais:

Composição, representação e modelagem de dados.

Aplicações Científicas

Workflow Científico

Além dos portais SINAPAD, outra abordagem que se destaca no LNCC são os **Workflows Científicos**. Um workflow científico consiste da especificação de um conjunto de aplicações científicas a serem executadas e suas dependências mútuas.

- Segue um ciclo de vida análogo ao dos experimentos científicos computacionais:

Composição, representação e modelagem de dados.

Mapeamento e execução.

Aplicações Científicas

Workflow Científico

Além dos portais SINAPAD, outra abordagem que se destaca no LNCC são os **Workflows Científicos**. Um workflow científico consiste da especificação de um conjunto de aplicações científicas a serem executadas e suas dependências mútuas.

- Segue um ciclo de vida análogo ao dos experimentos científicos computacionais:

Composição, representação e modelagem de dados.

Mapeamento e execução.

Coleta de metadados e proveniência.

Um sistema de gerência de workflows científicos (SGWC) permite gerenciar o ciclo de vida de workflows científicos. Contém um conjunto de técnicas e de ferramentas para a gerência de experimentos científicos computacionais compostos por muitas tarefas. O LNCC utiliza o SGWC **Swift** para implementação de workflows científicos.

Aplicações Científicas

Rodando uma aplicação em paralela no cluster

O usuário pode rodar uma aplicação paralela específica de seu interesse no cluster. Algumas das aplicações paralelas já instaladas e prontas para uso são:

BLAST (Basic Local Alignment Search Tool): aplicação em bioinformática que busca regiões de similaridade local entre as sequências biológicas

PETSc (Portable, Extensible Toolkit for Scientific Computation): um conjunto de pacotes de aplicação de amplo uso em ciência e engenharia. Resolução paralela de sistema de equações e de equações diferenciais para problemas de grande porte.

Openmodeller: pacote de algoritmos de análise de biodiversidade

Esta é apenas uma amostragem de aplicações. Em não havendo a aplicação paralela de interesse no cluster, o usuário poderá solicitar a instalação para os administradores.

Desenvolvendo uma aplicação paralela I

É possível que a aplicação de interesse do usuário ainda não esteja paralelizada, ou não possua todas as funcionalidades desejadas. Neste caso, o usuário que tiver algum conhecimento de programação, pode ele mesmo desenvolver módulos paralelos mediante o uso de bibliotecas já adaptadas para as arquiteturas multi-core ou many-core.

Bibliotecas de Álgebra Linear: gotoBLAS, ATLAS, MKL, scaLAPACK, MAGMA, ViennaCL, cuBLAS

Biblioteca de Otimização: C/C++ Minpack

Biblioteca de Dinâmica de Fluidos Computacional: OpenFOAM

Desenvolvendo uma aplicação paralela II

É possível ainda que as bibliotecas não atendam a uma demanda muito específica do usuário. Ou ele deseja paralelizar por si mesmo a sua aplicação. Nesta caso, o usuário que tiver bons conhecimentos de programação paralela, pode utilizar os compiladores e as ferramentas de avaliação de desempenho paralelo presentes no cluster.

- **Ferramentas de avaliação de desempenho:**

O primeiro passo em paralelizar um código, é saber onde ele gasta mais tempo processando

Para isso deve-se obter o **perfil de desempenho** do algoritmo sequencial

O perfil de desempenho mostra as rotinas que mais consomem processamento

São estas as rotinas a serem prioritariamente paralelizadas

Desenvolvendo uma aplicação paralela II

É possível ainda que as bibliotecas não atendam a uma demanda muito específica do usuário. Ou ele deseja paralelizar por si mesmo a sua aplicação. Nesta caso, o usuário que tiver bons conhecimentos de programação paralela, pode utilizar os compiladores e as ferramentas de avaliação de desempenho paralelo presentes no cluster.

- **Ferramentas de avaliação de desempenho:**

gprof: ferramenta open-source nativa de sistemas Linux. Uso por linha de comando.

Intel VTUNE: ferramenta da Intel. Além do tempo de cada rotina, extrai importantes informações quanto a possíveis limitantes de desempenho. Uso por linha de comando ou por interface gráfica.

LIKWID: open-source, também coleta informações importantes da execução, tais como o desempenho em GFlop/s. Uso por linha de comando.

Desenvolvendo uma aplicação paralela II (cont.)

PAPI (Performance Application Programming Interface): também open-source. Esta é uma biblioteca, para instrumentalizar o código da aplicação.

Desenvolvendo uma aplicação paralela III

Após a identificação dos chamados gargalos de desempenho, deve-se implementar a versão paralela do código. Há diversas opções de paralelizar o código. A escolha dependerá de fatores tais como do compilador e da arquitetura disponíveis, do conhecimento do programador e da adequação do código a estratégia escolhida.

MPI: biblioteca padrão de programação paralela em máquinas de memória distribuída. As implementações mais utilizadas são a MPICH, OpenMPI e MVAPICH.

OpenMP: padrão de programação paralela em máquinas de memória compartilhada, por meio de instrumentação de diretivas paralelas no código. A versão 4.0 da especificação foi lançada em Novembro de 2013, incluindo diretivas para uso de aceleradores.

OpenACC: padrão de programação paralela em aceleradores.

OpenCL: padrão de programação paralela **heterogênea**, ou seja, o mesmo código pode ser executado em diferentes dispositivos (CPU, GPU, FPGA, etc...).

Desenvolvendo uma aplicação paralela III (cont.)

CUDA: linguagem padrão para as GPUs NVIDIA.

Desenvolvendo uma aplicação paralela IV

- **Compiladores:**

GNU: `gcc`, `g++` e `gfortran`. Open-source e gratuito, sendo que, no entanto, geram normalmente códigos mais lentos que os compiladores comerciais.

Intel: `icc`, `icpc` e `ifort`. Em geral, obtém-se programas com maior desempenho, sobretudo em CPUs da própria Intel. Opções de otimização de código melhor adaptadas a arquitetura.

PGI (Portland Group/NVIDIA): `pgcc` e `pgfortran`. Compilador muito competitivo. Maior diferencial é a implementação do padrão OpenACC, para programação em aceleradores/coprocessadores (GPU e Xeon Phi).

Roteiro

1 Introdução

2 Recursos de *Hardware*

3 Recursos de *Software*

4 Considerações Finais

Considerações Finais

- O aluno de pós-graduação do LNCC pode fazer o pedido de conta nos clusters Alti-XE e SunHPC;

Considerações Finais

- O aluno de pós-graduação do LNCC pode fazer o pedido de conta nos clusters Alti-XE e SunHPC;
- Em qualquer etapa do mestrado ou doutorado;

Considerações Finais

- O aluno de pós-graduação do LNCC pode fazer o pedido de conta nos clusters Alti-XE e SunHPC;
- Em qualquer etapa do mestrado ou doutorado;
- Serão oferecidos cursos de extensão durante o ano;

Considerações Finais

- O aluno de pós-graduação do LNCC pode fazer o pedido de conta nos clusters Alti-XE e SunHPC;
- Em qualquer etapa do mestrado ou doutorado;
- Serão oferecidos cursos de extensão durante o ano;
- HPC pode ser praticado em casa (no seu laptop, desktop...);

Considerações Finais

- O aluno de pós-graduação do LNCC pode fazer o pedido de conta nos clusters Alti-XE e SunHPC;
- Em qualquer etapa do mestrado ou doutorado;
- Serão oferecidos cursos de extensão durante o ano;
- HPC pode ser praticado em casa (no seu laptop, desktop...);
- **Tendências:** arquiteturas mais eficientes quanto ao consumo de energia (ARM), Many-Core, ExaScale

Obrigado!