

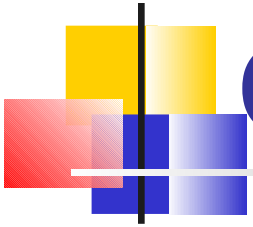
Exemplo de Arquitetura MIMD Clusters



SCE-200 – Arquiteturas Avançadas de Computadores

Profa. Dra. Regina H. C. Santana

ICMC-USP
Novembro/2005



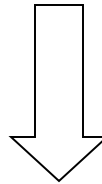
Conteúdo

1. Introdução
2. Conceitos Fundamentais
3. Ferramentas
4. Redes de Conexão
5. Análises e Comparações
6. Exemplos de Clusters
7. Conclusões

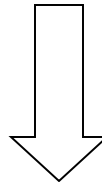


Introdução

Resolução de problemas científicos e comerciais
complexos de forma eficiente



Computação paralela



Supercomputadores e Clusters



Introdução

Desvantagens dos Supercomputadores:

- Utilização de softwares proprietários e caros;
- Alto custo de manutenção;
- Total dependência de fornecedores;
- Dificuldade de atualização;



Introdução

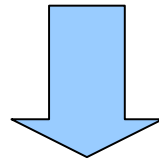
Cluster de Computadores

Clusters podem ser definidos como um grupo de computadores interconectados trabalhando de forma conjunta como um único computador. Os computadores do cluster se comportam como sistemas que podem, opcionalmente, funcionar de forma independente do cluster, mas que dão a impressão de uma imagem única ao usuário.



Introdução

Histórico: final de 1993, Donald Becker e Thomas Sterling (NASA) iniciaram um esboço de um sistema de processamento distribuído construído a partir de hardware convencional como uma medida de combate aos custos dos supercomputadores. No início de 1994, nasce o primeiro cluster e o projeto Beowulf.



Cluster de 16 processadores DX4 ligados por dois canais Ethernet acoplados.



Conceitos Fundamentais

- Grupo de computadores interconectados trabalhando de forma conjunta como um único computador
- Computadores do *cluster* (nós) se comportam como sistemas autônomos
- Normalmente, existe um nó mestre que gerencia e/ou divide as tarefas entre os demais nós, chamados de escravos.



Conceitos Fundamentais

Vantagens:

- Escalabilidade
- Tolerância a falhas
- Baixo Custo
- Independência de fornecedores



Conceitos Fundamentais

Tipos de Clusters:

- Alta Disponibilidade (base de dados, servidores de arquivos e aplicações);
- Balanceamento de carga (web farms);
- Combinação Alta Disponibilidade e Balanceamento de Carga;
- Processamento Paralelo ou Distribuído (divisão de tarefas);



Conceitos Fundamentais

Exemplos de Clusters:

➤ Beowulf Cluster

Desenvolvido na NASA com objetivo de utilizar computadores baratos e sistemas operacionais livres para montar *clusters*

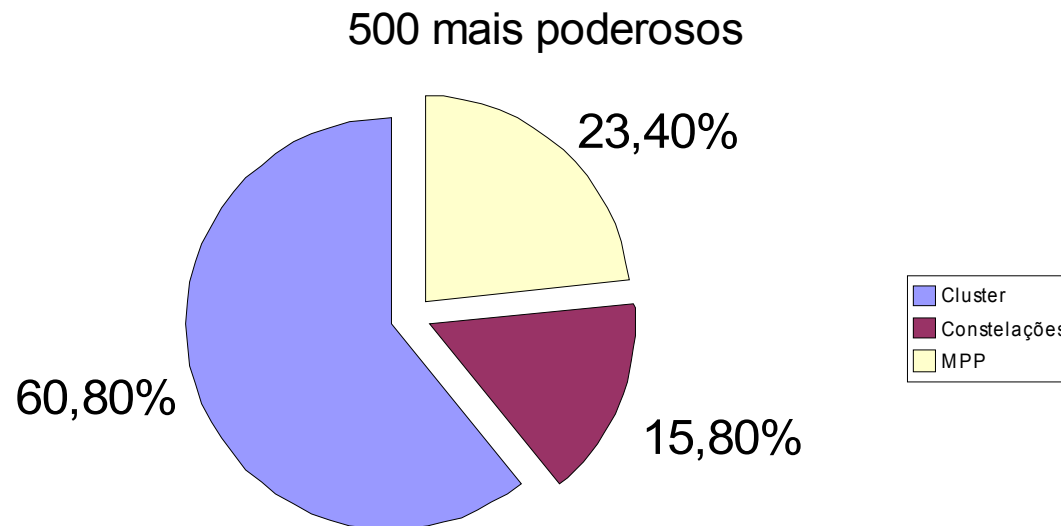
➤ Cluster OpenMosix

Extensão do projeto Mosix, baseado no próprio kernel do Linux



Conceitos Fundamentais

Dos 500 computadores mais poderosos do mundo, 304 são clusters





Ferramentas

O uso de ferramentas em *clusters* é importante para:

- Fornecer interface para o usuário utilizar o *cluster*
- Evitar sobrecarga nos nós
- Efetuar balanceamento de carga (*load balance*)



Ferramentas

Troca de Mensagens (*Message Passing*)

- Permite comunicação entre os nós o *cluster* através de uma rede de interconexão
- Independente de hardware, SO, linguagem de programação e bibliotecas.
- Bibliotecas de Comunicação: MPI e PVM



Ferramentas

MPI (*Message Passing Interface*)

- Padrão estabelecido pelo MPI Forum para troca de mensagens
- Especifica a sintaxe e semântica de um conjunto de funções de comunicação
- Modelo de programação SPMD (*Single Program Multiple Data*)
- Implementação em ANSI C e Fortran 77 (6 funções básicas e cerca de 125 funções no modo avançado)



Ferramentas

PVM (*Parallel Virtual Machine*)

- Software básico com suporte próprio ao gerenciamento de processos e centrado no conceito de máquina virtual.
- Modelo de programação MPMD (*Multiple Program Multiple Data*)
- Disponível em Java, Python, Perl, C, C++ e Fortran.
- Gerenciamento de recursos através da configuração de uma Máquina Virtual



Ferramentas

PVM (*Parallel Virtual Machine*)

- Possibilita a atribuição das sub-tarefas de uma aplicação, de forma otimizada, aos nós que compõem o ambiente paralelo;
- Apresenta uma interface de programação intuitiva e consistente;
- Oferece suporte para tolerância à falhas, monitoração e *profiling*;
- É altamente “portável”.



Ferramentas

MPI X PVM

<u>MPI</u>	<u>PVM</u>
Padrão Formal	Ainda não é Padrão Formal
SPMD (<i>Single Program Multiple Data</i>)	MPMD (<i>Multiple Program Multiple Data</i>)
Comunicação Coletiva e Ponto-a-Ponto eficiente	Comunicação Coletiva e Ponto-a-Ponto ruins
Equilíbrio entre facilidade de uso e desempenho para aplicações paralelas	Mecanismo de interoperabilidade, utilizando uma Máquina Virtual

Tendência de união das bibliotecas MPI e PVM.
Padrão MPI-2 – suporte ao MPMD.

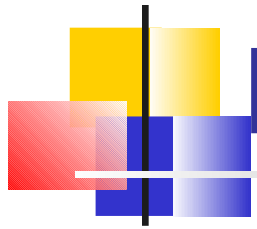


Redes de Conexão

Um dos gargalos mais sérios para o desempenho de *clusters* é a rede que interconecta seus computadores.

Redes de alto desempenho mais utilizadas em *clusters*:

- Myrinet
- InfiniBand
- Gigabit Ethernet



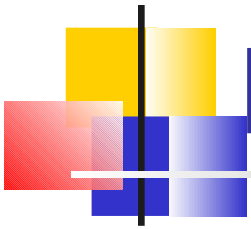
Redes de Interconexão

Gigabit Ethernet

É uma extensão dos padrões Ethernet e Fast Ethernet. Surgiu da necessidade criada pelo aumento de largura de banda nas “pontas” das redes.

Características:

- Tecnologia popular
- Baixo custo
- Não garante o fornecimento de QoS
- Pode atingir, atualmente, até 101 Gbps



Redes de Interconexão

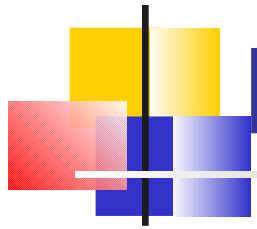
Myrinet

Rede baseada na tecnologia usada para comunicação e troca de pacotes entre processadores trabalhando em paralelo.

Implementa:

- Baixa latência
- Switches “*cut-through*”
- Altas taxas de transmissão
- Baixa taxa de erro
- Controle de fluxo em todos os links

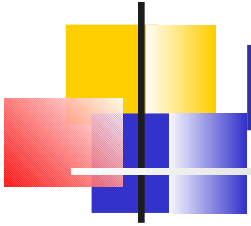
Permite comunicação direta entre processos do usuário e a rede.



Redes de Interconexão

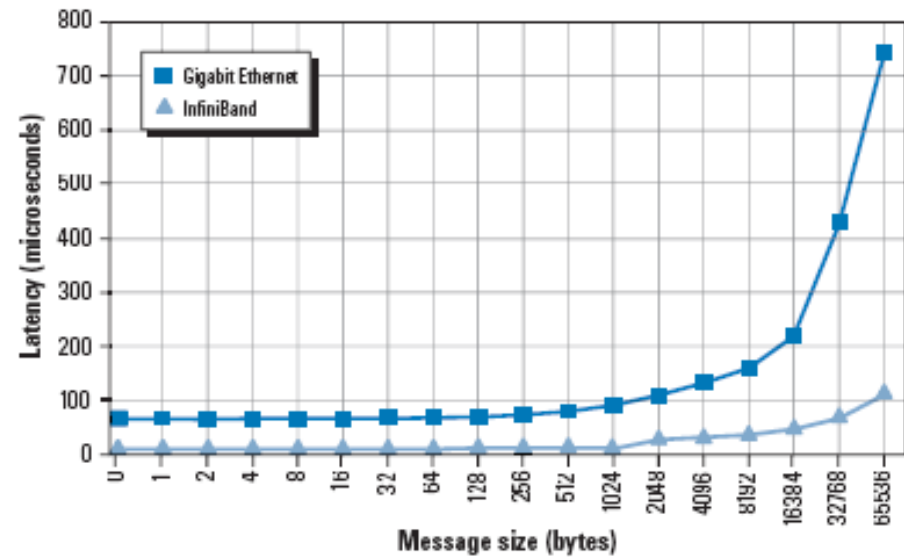
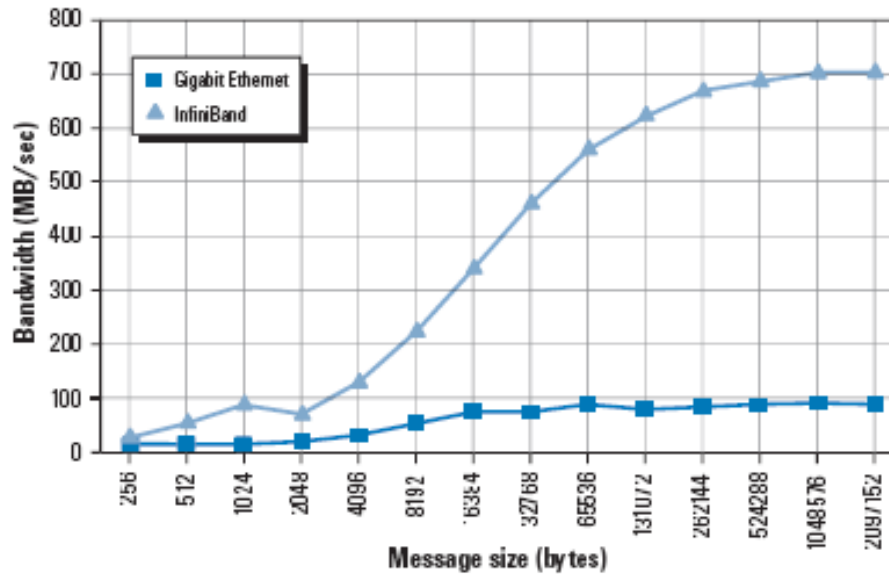
InfiniBand

- Surgiu da fusão de duas tecnologias concorrentes: Future I/O e Next Generation I/O, buscando atender à necessidade de se melhorar o desempenho dos dispositivos de E/S e da comunicação.
- É uma tecnologia adequada para aplicações HPC.
- Apresenta baixa latência e alta largura de banda
- Pretende-se incluir especificações de controle de congestionamento e QoS.



Redes de Interconexão

Comparação entre InfiniBand e Gigabit Ethernet





Análises e comparações

	<i>SMP</i>	<i>MPP</i>	<i>NOW</i>
Número de Processadores	baixo	alto	médio
Escalabilidade	baixa	alta	média
Latência da Comunicação	média	baixa	alta
Programação	fácil	difícil	difícil
Custo	médio	alto	baixo

SMP - *Symetric Multiprocessors*

MPP - *Massively Parallel Processors*

NOW-*Network Of Workstations* (Cluster)



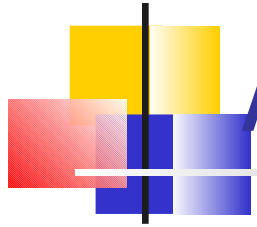
Análises e comparações

Clusters

- Comunicação via bus de I/O
- Conexão baseada em software
- Alta escalabilidade
- Alto custo de manutenção
- Alta disponibilidade
- Melhor custo/benefício

Máquinas Multi-processadas com memória compartilhada

- Comunicação via bus de memória
- Conexão baseada em hardware
- Baixa escalabilidade
- Custo mais baixo de manutenção



Análises e comparações

Desvantagens dos Clusters:

- Difícil Programação;
- Altos Custos de Manutenção e Administração;
- Alta latência de comunicação;
- Utilização do bus de I/O;



Exemplos

Cluster do sistema de busca do Google:

- Cluster de Load Balancing
- 15.000 PCs
- Conectado por Fast Ethernet e Gigabit Ethernet
- Menor custo com energia



Exemplos

Cluster

88 dual-CPU 2GHz Intel Xeon
2 Gbytes de RAM
disco de 80 GB

\$278,000

Capacidade resultante:
176 processadores de 2GHz
176 Gbytes de RAM
7 Tbytes de disco

Máquina Multiprocessada

8 processadores 2-GHz Xeon
64 Gbytes de RAM
Disco de 8 Tbytes

\$758,000

3 vezes mais caro
22 vezes menos CPUs
3 vezes menos memória



Exemplos

MareNostrum - JS20 Cluster

- Barcelona Supercomputer Center;
- 5º mais poderoso;
- 4800 processadores (2406 nós dual)
PowerPC 2200 MHz (8.8 GFlops);
- Pico de processamento: 42144 GFlops;
- Sistema Operacional Linux;



Exemplos

MareNostrum - JS20 Cluster

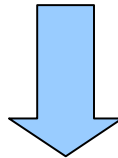
- Rede de Interconexão: Myrinet;
- 42 racks, ocupando 120 m²;
- Parceria entre o governo espanhol e a IBM;
- Fins científicos;



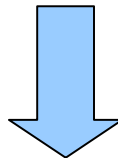
Exemplos

No Brasil:

- 4 computadores na lista dos 500 mais poderosos do mundo



Todos clusters !!



Todos da Petrobrás !!



Exemplos

xSeries Cluster Xeon 3.06 GHz – GigE

- 95º posição;
- 1024 processadores Intel Pentium 4 Xeon 3060 MHz (6.12 Gflops);
- Pico de processamento: 6288.88 GFlops;
- Rede de Interconexão: Gigabit Ethernet;
- Sistema Operacional Linux;
- Área de Aplicação: Geofísica;



Exemplos

xSeries Cluster Xeon 3.06 GHz – GigE

Além deste cluster, Petrobrás possui outros 3 clusters na lista dos 500 mais poderosos, ocupando a 97º, 118º e 208º posição.



Exemplos

Na USP:

- Caverna Digital;

- Cluster do ICMC da USP em São Carlos: utilizado para processamento científico e matemático.
 - 12 nós de processamento, cada um com 2 processadores Intel® Xeon™ 2.40GHz, totalizando 24 processadores;



Exemplos

Na USP:

- Projeto para ter um cluster, que entrará na lista dos 500 mais poderosos;
- Cluster Sun;
- 624 chips "dual core" da AMD e S.O. Linux;
- US\$ 650 mil;
- Capacidade de Processamento: 2,8 teraflops (301º posição);



Conclusão

Principais vantagens de clusters:

- Cluster são uma boa alternativa para uso de processamento de alto desempenho a um custo mais acessível.
- Boa Escalabilidade, devido à facilidade de adição de novas máquinas.
- Alta disponibilidade
- Novas tendências tecnológicas em hardware e software permitirão aos *clusters* ter o desempenho de um supercomputador.