

A Evolução dos Clusters frente as demais arquiteturas de Alto Desempenho

Rodrigo Santos de Souza, Adenauer C. Yamin

Universidade Católica de Pelotas - UCPel
{rsouza,adenauer}@ucpel.tche.br,

1 Introdução

As arquiteturas utilizadas nos supercomputadores são bastante variadas nas suas diferentes particularidades, como são máquinas especiais esses equipamentos por muitos anos costumavam fugir dos padrões comerciais de hardware, mas algumas mudanças estão sendo observadas neste sentido. Os custos para se construir uma máquina destas é bastante elevado e por isso nos últimos anos é notória a ascensão da utilização de clusters em aplicações de alto desempenho. O uso de hardwares padronizados na constituição dos nodos ao invés de peças feitas sob encomenda reduz bastante o custo final do equipamento.

O objetivo deste trabalho é caracterizar esta tendência categorizando as arquiteturas que aparecem na lista dos 500 supercomputadores mais poderosos do mundo (www.top500.org) disponibilizando estas informações de forma gráfica para que possa facilmente ser consultada e utilizada por profissionais da área.

2 Tipos de arquiteturas paralelas

O objetivo neste momento é introduzir uma breve explicação sobre os tipos de arquiteturas paralelas que serão usados no texto, assim como nas tabelas e gráficos[2][1].

2.1 SIMD (Single Instruction stream, Multiple Data stream):

Modelo de execução em paralelo no qual todos os processadores executam a mesma operação no mesmo instante de tempo, porém cada um opera com seu próprio conjunto de dados. Este tipo de estrutura é muito adequada para cálculos complexo com grandes matrizes e vetores, como são o caso em algumas áreas da física, da meteorologia, engenharia, entre outras. Devido ao tipo de aplicação associada, esses computadores são chamados computadores vetoriais.

2.2 MIMD (Multiple Instruction stream, Multiple Data stream):

Neste modelo todos os processadores agem independentemente uns dos outros, permitindo que um programa possa ser decomposto em diferentes partes para execução em paralelo. Este modelo é mais flexível que o SIMD.

2.3 SMP (Symmetric Multi-Processor):

A arquitetura SMP é uma arquitetura do tipo MIMD que usa memória compartilhada. É composta por um grupo de processadores com as mesmas propriedades trabalhando em conjunto e que utilizam todos a mesma memória física, de forma que qualquer um têm o mesmo desempenho se executar o mesmo trecho de código.

2.4 Constellation Cluster

Constellation Cluster consiste de um agregado de multiprocessadores.

2.5 MPP (Massively Parallel Processor):

Sistema composto de um grande número de processadores (as vezes até milhares), cada qual com sua respectiva memória, disco rígido e sistema operacional, utilizado para a solução de grandes problemas em computação.

2.6 Cluster

Cluster pode ser visto como um tipo particular de sistema MPP onde os nodos são tipicamente computadores de padrão comercial, e o número de nodos é menor.

3 A evolução das arquiteturas paralelas

O site top500 é referência para a obtenção informações sobre os computadores de maior desempenho no mundo. Através do mesmo é possível ter uma visão do futuro das arquiteturas e as tendências nesta área. Baseado nisto, foram feitas consultas em seu banco de dados com o intuito de observar a evolução dos tipos de arquiteturas usadas nestas máquinas nos últimos 10 anos. A cada ano são disponibilizadas duas listas, uma em junho e outra em novembro, cada uma contendo 500 supercomputadores segundo uma ordem baseada no poder de processamento[3]. Para este trabalho foi levado em conta apenas as listas dos meses de novembro. A totalidade de cada lista foi tomada como base, tomando-se apenas os percentuais de cada arquitetura sem preocupação com o número exato de máquinas de cada tipo.

A Tabela 1 lista o percentual de máquinas de cada arquitetura.

Como o intuito do trabalho é avaliar a evolução das arquiteturas tipo cluster, foi feita uma segunda tabela em que as ocorrências das demais arquiteturas são somadas, para que sejam observadas em conjunto perante a evolução dos clusters. A Tabela 2 traz esta comparação.

A Tabela 1 deu origem ao gráfico da Figura 1, assim como a Tabela 2 o Gráfico 2.

Pela observação, principalmente da Figura 2, fica claro a evolução dos cluster entre as arquiteturas de alto desempenho, sendo que no último ano a predominância desta arquitetura é massiva, chegando a 72,2% do total das máquinas da lista.

	Constellations	MPP	Cluster	SMP	SIMD	Sing. Processor
11/96	3,8	57,6	0	36,6	1,4	0,6
11/97	2	45,2	0,2	52,6	0	0
11/98	3,4	45,2	0,4	51	0	0
11/99	13,2	51,6	1,4	33,8	0	0
11/00	23	69,2	5,6	2,2	0	0
11/01	28,6	51,4	8,6	11,4	0	0
11/02	40,8	40,6	18,6	0	0	0
11/03	25,4	32,6	42	0	0	0
11/04	21,4	19,8	58,8	0	0	0
11/05	7,2	20,6	72,2	0	0	0
11/06	6,2	21,6	72,2	0	0	0

Tabela 1: TIPOS DE ARQUITETURAS - TOP500

4 Considerações finais

O objetivo do trabalho é caracterizar o crescimento do número de estruturas de tipo cluster e sua atual predominância numérica em relação as demais arquiteturas. A fonte para isto são os dados pertinentes aos equipamentos reais disponíveis nos bancos de dados do site top500.

Referências

- [1] Leonardo de Souza Miers. *Implementação do Método dos Elementos de Contorno para Eletricidade Tridimensional em Ambiente Paralelo de Memória Distribuída*. PhD thesis, Universidade Federal do Rio de Janeiro, 2003.
- [2] Johnny Marcus Gomes Rocha. *Cluster beowulf: Aspectos de projeto e implementação*. Master's thesis, Universidade Federal do Paraná, 2003.
- [3] Top500. www.top500.org, Abril 2007.

	Cluster	Outros
11/96	0	100
11/97	0,2	99,8
11/98	0,4	99,6
11/99	1,4	98,6
11/00	5,6	94,4
11/01	8,6	91,4
11/02	18,6	81,4
11/03	42	58
11/04	58,8	41,2
11/05	72,2	27,8
11/06	72,2	27,8

Tabela 2: CLUSTERS VERSUS DEMAIS ARQUITETURAS - TOP500

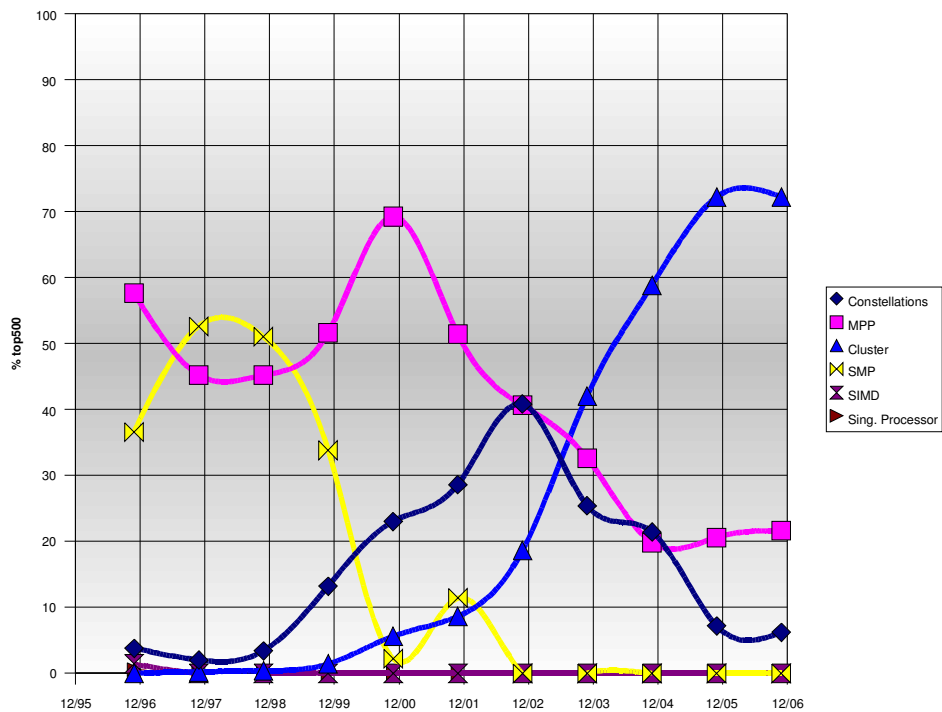


Figura 1: CURVA DE EVOLUÇÃO DOS TIPOS DE ARQUITETURAS - TOP500

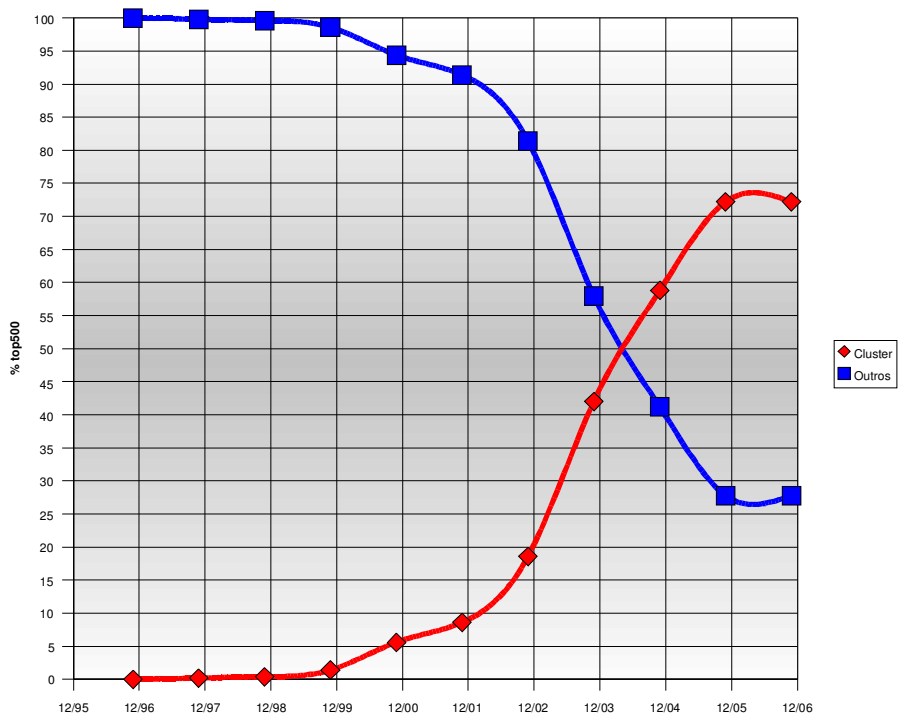


Figura 2: CURVA DE EVOLUÇÃO DOS CLUSTERS FRENTE AOS DEMAIS - TOP500