

Introdução às arquiteturas paralelas e taxonomia de Flynn

OBJETIVO: DEFINIR COMPUTAÇÃO PARALELA; O MODELO DE COMPUTAÇÃO PARALELA DESEMPENHADA POR COMPUTADORES PARALELOS; E EXEMPLOS DE USO DA ARQUITETURA PARALELA.

Introdução

O grande interesse por problemas cada vez mais complexos tem levado à necessidade de computadores cada vez mais potentes para resolvê-los.

Entretanto, limitações físicas e econômicas têm restringido o aumento da velocidade dos computadores sequenciais, ou seja, computadores que executam instruções em série, uma após a outra pela CPU. Por outro lado, os problemas computacionais usualmente podem ter algumas de suas partes divididas em pedaços que teriam como ser solucionados ao mesmo tempo ou processadas em paralelo. Processamento paralelo é então uma forma pela qual a demanda computacional é suprida por meio do uso simultâneo de recursos computacionais como processadores para solução de um problema.

A computação paralela é caracterizada pelo uso de várias unidades de processamento ou processadores para executar uma computação de forma mais rápida. É baseada no fato de que o processo de resolução de um problema pode ser dividido em tarefas menores, realizadas simultaneamente por meio de algum tipo de coordenação. O conceito foi originalmente introduzido no CDC 6600 em 1964 pela CDC (control data corporation). No tópico a seguir, serão descritos os modelos de computação paralela existentes.

Taxonomia de Flynn

A taxonomia de Flynn abrange quatro classes de arquiteturas de computadores:

- SISD (*Single Instruction Single Data*): fluxo único de instruções sobre um único conjunto de dados.
- SIMD (*Single Instruction Multiple Data*): fluxo único de instruções em múltiplos conjuntos de dados.
- MISD (*Multiple Instruction Single Data*): fluxo múltiplo de instruções em um único conjunto de dados.
- MIMD (*Multiple Instruction Multiple Data*): fluxo múltiplo de instruções sobre múltiplos conjuntos de dados.

SISD (*Single Instruction Single Data*)

Nesta classe, um único fluxo de instruções opera sobre um único fluxo de dados. Isso corresponde ao processamento sequencial característico da máquina de Von Neumann e que compreende os computadores pessoais e estações de trabalho. Apesar dos programas estarem organizados por meio de instruções sequenciais, elas podem ser executadas de forma sobreposta em diferentes estágios (*pipelining*). Arquiteturas SISD caracterizam-se por possuírem uma única unidade de controle, podendo possuir mais de uma unidade funcional.

Um exemplo seria seu computador pessoal com um processador convencional.

SIMD (Single Instruction Stream Multiple Data)

Esta classificação corresponde ao processamento de vários dados sob o comando de apenas uma instrução. Em uma arquitetura SIMD, o programa ainda segue uma organização sequencial. Para possibilitar o acesso a múltiplos dados é preciso uma organização de memória em diversos módulos.

A unidade de controle é única, e existem diversas unidades funcionais. Nesta classe estão os processadores vetoriais e matriciais.

Quanto às facilidades de hardware para armazenamento, essas normalmente são classificadas como:

- Processor Array
- Vector Pipeline

Exemplo de computadores com a arquitetura processor array são as máquinas ILLIAC IV (Universidade de Illinois), Thinking Machine CM- 2 e MASPAR MP-1216.

Exemplo de computadores com a arquitetura vector pipeline são as máquinas IBM 9000, Cray X-MP, Y-MP & C90, Fujitsu VP, NEC SX-2, Hitachi S820, ETA10. As GPUs, também, estão sob essa classificação.

MISD (Multiple Instruction Stream Single Data Stream)

Um conjunto de dados é colocado concorrente em múltiplas unidades de processamento. Cada UP opera de maneira independente via conjuntos independentes de instruções.

Algumas utilizações de uma configuração MISD poderiam ser:

- Filtros de múltiplas frequências operando um mesmo sinal
- Múltiplos algoritmos de criptografia tentando a quebra de uma mensagem codificada

Não se tem conhecimento de arquitetura de máquinas comercial com múltiplas instruções trabalhando com um único conjunto de dados concorrente. Em 1971, uma máquina denominada como C.mmp computer foi desenvolvida na Universidade de Carnegie-Mellon.

MIMD (Multiple Instruction Multiple Data)

Esta classe é bastante genérica, envolvendo o processamento de múltiplos dados por parte de múltiplas instruções. Neste caso, várias unidades de controle comandam suas unidades funcionais, as quais têm acesso a vários módulos de memória. Qualquer grupo de máquinas operando como uma unidade (deve haver certo grau de interação entre as máquinas) enquadra-se como MIMD. Alguns representantes desta categoria são os servidores multiprocessados, as redes de estações e as arquiteturas massivamente paralelas.

Aplicação do uso da arquitetura paralela

Processadores vetoriais ou processamento vetorial é definido como aplicação de operações aritméticas/lógicas sobre vetores, em vez de operações sobre pares de dados ou dados escalares.

Ela possui as seguintes características: reduz o custo de manutenção de estruturas de controle para laços de processamento; reduz conflitos de acesso à memória; pode ser empregada com conceitos de pipeline; em geral tem um aumento de desempenho da ordem de 10 a 20 vezes quando comparada com máquinas escalares; aumenta os custos de hardware; aumenta os custos de compilação do código (vetorização).

Multiprocessadores simétricos

Estes ambientes são conhecidos como arquiteturas de compartilhamento total, que são caracterizadas por até dezenas de processadores compartilhando os mesmos recursos computacionais e rodando um único sistema operacional. Os processadores são considerados simétricos porque têm os mesmos custos para acesso à memória principal.

A utilização de SMP é mais popular do que se imagina. Esse tipo de máquina é encontrado facilmente em grande parte das organizações de hoje, e também vem ganhando espaço em áreas menores, reflexo da redução de custos.

Um problema desta arquitetura é sua escalabilidade, pois, com o aumento do número de processadores, a taxa de colisão de acesso à memória também cresce, sendo necessária a utilização de soluções de memórias de cache e globais, que serão vistas à frente.

Máquinas maciçamente paralelas

Como o próprio nome já diz, essas máquinas almejam o alto desempenho por meio da utilização de um grande número de processadores comerciais, os quais, por causa do fator do custo, acabam sendo processadores de baixo ou médio poder computacional.

Cada nó possui uma memória local com um espaço de endereçamento próprio. Assim, cada nó só tem acesso à sua própria memória, o que a caracteriza como uma máquina NORMA. Assim, a comunicação entre as máquinas é feita por meio de troca de mensagens.

Máquinas com memória compartilhada distribuída

São sistemas em que, apesar das memórias estarem fisicamente separadas, todos os processadores podem endereçar as memórias de todos os nós por causa do fato de que foi implementado um espaço único de endereçamento (que pode ter sido feito tanto em hardware como em software).

Rede de estações de trabalho

Basicamente as máquinas NOW utilizam uma rede local (normalmente Ethernet ou ATM) já existente para a execução de aplicações paralelas. A rede local pode ser vista como uma máquina paralela em que vários processadores, com suas memórias locais (estações de trabalho), são interligados por uma rede, o que a torna uma máquina NORMA de custo quase nulo.

A diferença da NOW para as MPP está na hierarquia de barramento utilizada nas estações, por possuir um disco local (DL) nos nós e na rede de interconexão. Essas diferenças vêm do fato de a rede local ter sido planejada para um tipo distinto de aplicações paralelas (compartilhar arquivos e acessar periféricos remotos, como a impressora).

Uma das dificuldades encontradas nessa troca de função está na rede de interconexão. Redes como Ethernet e ATM não são otimizadas para aplicações paralelas, que geram uma alta latência nas operações, o que compromete o desempenho da máquina como um todo.

Cluster

Um cluster, ou aglomerado de computadores, é formado por um conjunto de computadores que se utiliza de um tipo especial de sistema operacional classificado como sistema distribuído. Muitas vezes, é construído a partir de computadores convencionais (personal computers), os quais são ligados em rede e comunicam-se por meio do sistema, trabalhando como se fossem uma única máquina de grande porte. Há diversos tipos de cluster. Um tipo famoso é o cluster da classe Beowulf, constituído por diversos nós escravos gerenciados por um só computador.

Ele pode ser definido também como o cruzamento de uma trilha com um setor de um disco formatado. Um HD (hard disc) possui vários clusters que serão utilizados para armazenamento dos dados de um determinado arquivo. Com essa divisão em trilhas e setores, é possível criar um endereçamento que visa facilitar o acesso a dados não contíguos, assim como o endereçamento de uma planilha de cálculos ou, como um exemplo mais simples, o tabuleiro do jogo "batalha naval".

REFERÊNCIA

STALLINGS, Willian. Arquitetura e organização de computadores. 5. ed. Prentice Hall. São Paulo, 2006.

TANENBAUM. Andrew S. Organização estruturada de computadores. 5. ed. Rio de Janeiro: LTC, 2007.