

Arquitetura de Computadores

Prof. Dr. José Augusto de Sena Quaresma Jq.quaresma12@gmail.com

Revisão simulado avaliando aprendizagem

- Um computador pra funcionar é dividido em duas camadas fundamentais. Quais são essas duas camadas fundamentais do computador?
 - CPU e memória.
 - Software e usuário.
 - Aplicativo e browser.
 - Kernel e interface.
 - Hardware e software.

- Sobre os dispositivos de entrada de dados é correto afirmar que:
 - São dispositivos que o computador pode funcionar normalmente sem nenhum eles, não fazendo diferença a existência ou não deles no sistema computacional.
 - São exemplos de dispositivos de entrada de dados: plotter, caixa de som, impressora, teclado.
 - São dispositivos para manter os dados gravados de forma permanente em um meio físico permitindo sua leitura futura.
 - É todo e qualquer dispositivo que compõe o computador, tendo como finalidade o envio dos dados do mesmo (como informações úteis) ao mundo exterior.
 - São todos os dispositivos que fazem parte do sistema computacional com a finalidade realizar a entrada de dados no computador (sejam comandos, dados ou outras informações externas ao computador).

- Considerando o valor 959 na base 10, qual é o valor correspondente na base 2?
 - > 001100111111
 - > 001110111111
 - > 011010111111
 - > 000010111111
 - > 001010111111

```
959 / 2 = 479 (resto 1)

479 / 2 = 239 (resto 1)

239 / 2 = 119 (resto 1)

119 / 2 = 59 (resto 1)

59 / 2 = 29 (resto 1)

29 / 2 = 14 (resto 1)

14 / 2 = 7 (resto 0)

7 / 2 = 3 (resto 1)

3 / 2 = 1 (resto 1)

1 / 2 = 0 (resto 1)
```

- Assinale a alternativa que apresenta a denominação utiliza da para "um grupo ordenado de 8 bits, tratados de forma individual, como unidade de armazename nto e transferência".
 - > Bit
 - Palavra
 - Byte
 - KiloByte
 - Caractere

- Os múltiplos da base 2 são valores que representam potências de 2. Eles são frequentemente usados na computação e na representação de unidades de informação, como bits e bytes. Qual é o prefixo de múltiplos usado para representar 1.024 megabytes?
 - Quilo
 - Giga
 - Mega
 - > Tera
 - > Kilo

- Durante um seminário sobre arquitetura de computadores, um especialista enfatizou o papel dos bits na representação de dados. Ele explicou como um único bit pode representar um estado binário, sendo a base para estruturas de dados mais complexas. O que um único bit pode representar em um sistema de computação?
 - Um código de cor em um pixel.
 - Um valor binário (0 ou 1).
 - Um valor decimal.
 - Uma sequência de caracteres.
 - Um número inteiro.

- A conversão de bases numéricas é um processo fundamental na computação, pois os sistemas de computadores usam principalmente a base binária (base 2) para representar dados. No entanto, em determinadas situações, é necessário converter entre diferentes bases numéricas, como base 2 (binária), base 8 (octal), base 10 (decimal) e base 16 (hexadecimal). Qual é o equivalente em hexadecimal do número binário 101010?
 - > C6
 - > B4
 - > EA
 - > 2A
 - ▶ D8

➤ A compreensão da conversão entre bases numéricas é crucial para programadores, especialmente ao lidar com programação de baixo nível, sistemas embarcados e manipulação de dados binários. Qual é o valor em decimal do número binário 1011?

- **>** 14
- **≻** 12
- **>** 13
- **>** 11
- **>** 10

- Em um workshop sobre armazenamento de dados, o instrutor explicou diferentes unidades de medida para dados digitais. Começando pelo bit, ele detalhou como essas unidades são escaladas para representar quantidades maiores de dados em sistemas de computação. Qual das seguintes é a maior unidade de medida de dados?
 - Gigabyte
 - Megabyte
 - Byte
 - Kilobyte
 - Terabyte

- Considere a operação 17 + 10 = X em base numérica decimal e marque a alternativa que mostra o resultado correto de X em uma base Hexadecimal.
 - **>** 17
 - > 1B
 - > B1
 - **>** 27
 - > 2B

Processamento Paralelo

Processamento paralelo

- Conceito: é uma técnica de processamento de informações na qual várias tarefas são executadas simultaneamente. Em vez de executar uma tarefa após a outra, o processamento em paralelo permite que várias tarefas sejam realizadas ao mesmo tempo, aumentando assim a eficiência e a velocidade do processamento.
- Amplamente utilizado em uma variedade de campos, incluindo computação de alto desempenho, inteligência artificial, processamento de dados em tempo real, renderização de gráficos, simulações científicas e muito mais. Ele oferece vantagens significativas em termos de desempenho e capacidade de lidar com cargas de trabalho intensivas.

Abordagens

- Processamento paralelo em hardware
- Processamento paralelo em software
- Computação distribuída

Processamento paralelo em hardware

- Refere-se à capacidade de realizar múltiplas operações ou tarefas simultaneamente por meio de dispositivos ou componentes de hardware. Esta abordagem visa aumentar a eficiência e o desempenho do sistema, permitindo que várias operações sejam executadas em paralelo.
- São essenciais para lidar com cargas de trabalho cada vez mais exigentes em termos de computação e para alcançar um desempenho escalável em sistemas modernos.

Implementação

- Multiprocessamento Simétrico (SMP)
- Arquiteturas de Processadores Multicore
- Arquiteturas de Processadores SIMD e Vector Processing
- Graphics Processing Unit

Multiprocessamento Simétrico (SMP)

- Abordagem de processamento paralelo em hardware que envolve a utilização de múltiplos processadores ou núcleos de processamento em um único sistema. Nesse modelo, todos os processadores compartilham o acesso uniforme à memória principal e têm a capacidade de executar tarefas simultaneamente.
- Comumente encontrado em uma variedade de sistemas, desde servidores de alto desempenho até desktops e laptops modernos, nos quais vários núcleos de processamento são usados para lidar com cargas de trabalho exigentes e executar tarefas simultaneamente.

Características-chave do SMP

- Acesso Uniforme à Memória: Todos os processadores em um sistema SMP têm acesso uniforme à memória principal. Isso significa que qualquer processador pode acessar qualquer endereço de memória sem restrições, proporcionando uma visão consistente da memória compartilhada.
- Escalabilidade Simétrica: O SMP permite a adição de processadores adicionais ao sistema de forma simétrica, ou seja, todos os processadores têm o mesmo acesso e autoridade sobre os recursos do sistema. Isso proporciona uma escalabilidade relativamente simples, pois novos processadores podem ser adicionados sem alterações significativas na arquitetura do sistema.

Características-chave do SMP (2)

- Sistemas de Multiprocessamento de Núcleo: Com o avanço da tecnologia, muitos sistemas modernos adotam o conceito de multiprocessamento de núcleo, nos quais múltiplos núcleos de processamento são integrados em um único chip. Esses sistemas ainda seguem os princípios do SMP, permitindo que todos os núcleos compartilhem a mesma memória e executem tarefas de forma simultânea.
- Divisão de Tarefas: Em um sistema SMP, as tarefas podem ser divididas entre os diferentes processadores de forma eficiente, permitindo que múltiplas tarefas sejam executadas em paralelo. Isso pode resultar em uma melhoria significativa no desempenho e na capacidade de resposta do sistema, especialmente em aplicações que podem ser paralelizadas.

Arquitetura de Processadores Multicore

- Refere-se a um tipo de design de CPU que integra múltiplos núcleos de processamento em um único chip. Cada núcleo é capaz de executar instruções de forma independente, permitindo que múltiplas tarefas sejam processadas simultaneamente.
- Papel fundamental no fornecimento de desempenho escalável e eficiente para uma ampla gama de aplicações computacionais

Arquitetura de Processadores Multicore – Características

- Núcleos de Processamento: Um processador multicore possui dois ou mais núcleos de processamento em um único chip. Cada núcleo funciona como uma unidade de processamento independente, capaz de executar instruções e acessar memória de forma autônoma.
- Compartilhamento de Recursos: Apesar de cada núcleo ser capaz de executar tarefas independentes, eles compartilham recursos como cache de memória, controladores de memória e barramentos de dados. Isso permite uma utilização eficiente dos recursos do sistema e ajuda a minimizar gargalos de comunicação entre os núcleos.

Arquitetura de Processadores Multicore – Características

- Escalabilidade: A arquitetura multicore oferece uma forma escalável de aumentar o poder de processamento de um sistema sem a necessidade de aumentar a frequência do clock do processador. Adicionando mais núcleos, o sistema pode lidar com uma carga de trabalho maior e executar mais tarefas simultaneamente.
- Paralelismo de Tarefas: Com múltiplos núcleos de processamento, as tarefas podem ser distribuídas entre os núcleos de forma a maximizar o paralelismo. Isso é especialmente útil em aplicações que podem ser divididas em tarefas independentes, como processamento de vídeo, renderização de gráficos, simulações científicas e computação de alto desempenho.

Arquitetura de Processadores Multicore – Características

Fficiência Energética: Comparado com aumentar a frequência do clock do processador para melhorar o desempenho, o uso de múltiplos núcleos pode ser mais eficiente em termos de energia. Isso permite uma maior eficiência energética em sistemas que requerem alto desempenho computacional, como servidores e dispositivos móveis.

Arquiteturas de Processadores SIMD e Vector Processing

- Projetadas para processar múltiplos elementos de dados simultaneamente, utilizando uma única instrução.
- Essas arquiteturas são especialmente eficientes para lidar com operações repetitivas sobre grandes conjuntos de dados, como operações matriciais, processamento de imagens, processamento de sinais digitais e simulações científicas.

Arquiteturas de Processadores SIMD e Vector Processing – Características

- Instruções SIMD: Em uma arquitetura SIMD, uma única instrução é aplicada a múltiplos elementos de dados ao mesmo tempo. Por exemplo, uma instrução SIMD pode somar, subtrair ou multiplicar simultaneamente elementos correspondentes em dois vetores.
- Registradores SIMD: Os processadores SIMD contêm registradores especiais de grande largura, capazes de armazenar múltiplos elementos de dados em paralelo. Cada operação SIMD é realizada nos dados armazenados nesses registradores, o que permite um processamento eficiente de grandes volumes de dados.
- Paralelismo de Dados: O paralelismo de dados é explorado ao processar múltiplos elementos de dados simultaneamente. Isso pode resultar em um desempenho significativamente maior em comparação com arquiteturas que processam apenas um elemento de dados por vez.

Arquiteturas de Processadores SIMD e Vector Processing – Características – 2

- Aplicações: As arquiteturas SIMD são amplamente utilizadas em uma variedade de aplicações, incluindo processamento de vídeo e áudio, jogos, criptografia, compressão de dados, simulações científicas e muito mais. Qualquer aplicação que envolva operações intensivas em dados pode se beneficiar do uso de instruções SIMD.
- Extensões SIMD: Muitos conjuntos de instruções de arquiteturas de processadores modernos incluem extensões SIMD dedicadas, como o SSE (Streaming SIMD Extensions) da Intel e o NEON da ARM. Essas extensões fornecem suporte adicional para operações SIMD e são amplamente utilizadas no desenvolvimento de software otimizado para desempenho.
- Vector Processing: O vector processing é uma forma de processamento de dados que envolve a execução de operações em vetores ou matrizes de dados. Embora muitas vezes associado às arquiteturas SIMD, o vector processing pode ser implementado de várias maneiras, incluindo unidades de processamento vetorial dedicadas e operações de vetorização em hardware convencional.

Graphics Processing Unit

- Um tipo de processador especializado projetado para lidar com operações relacionadas a gráficos e processamento paralelo em grande escala.
- Enquanto as CPUs (Unidades de Processamento Central) são otimizadas para tarefas de propósito geral, como processamento de texto e cálculos matemáticos, as GPUs são otimizadas para lidar com cálculos intensivos em paralelo, como renderização de gráficos, simulações científicas, aprendizado de máquina e mineração de criptomoedas.

Graphics Processing Unit – Características

- Arquitetura Paralela: As GPUs são compostas por milhares de núcleos de processamento (também conhecidos como shaders ou CUDA cores), organizados em centenas de unidades de processamento chamadas de Streaming Multiprocessors (SMs). Essa arquitetura paralela permite que a GPU processe várias tarefas de forma simultânea, resultando em um alto desempenho computacional.
- Processamento Massivamente Paralelo: As GPUs são altamente eficientes em lidar com tarefas que podem ser divididas em pequenas partes e processadas em paralelo. Isso é especialmente útil em aplicações gráficas, onde milhões de pixels podem ser processados simultaneamente para renderizar imagens complexas em tempo real.

Graphics Processing Unit – Características 2

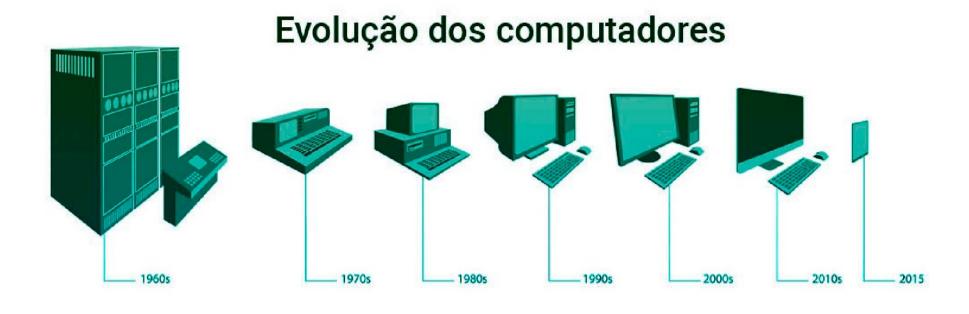
- Memória Rápida e Largura de Banda Elevada: As GPUs possuem memória rápida e largura de banda elevada para permitir a transferência eficiente de dados entre o processador e a memória. Isso é crucial para lidar com conjuntos de dados grandes, como texturas, modelos 3D e dados de treinamento em aprendizado de máquina.
- ▶ GPGPU (General-Purpose Computing on Graphics Processing Units): Além de processar gráficos, as GPUs são cada vez mais utilizadas para computação de propósito geral. Com frameworks como CUDA (NVIDIA) e OpenCL, os desenvolvedores podem escrever programas que aproveitam o poder de processamento das GPUs para uma variedade de aplicações, incluindo simulações científicas, análise de dados, criptografia e muito mais.

Graphics Processing Unit – Características 3

Aplicações Diversificadas: As GPUs são amplamente utilizadas em uma variedade de campos, incluindo jogos, design gráfico, animação, modelagem 3D, renderização de filmes, ciência da computação, aprendizado de máquina, finanças, ciências médicas e muito mais. Sua capacidade de processamento massivamente paralelo torna-as uma ferramenta essencial para acelerar uma ampla gama de tarefas computacionais exigentes.



Evolução dos computadores



A evolução da máquina.

Limitações de processamento



Processamento de imagem



Renderização tridimensional



Reconhecimento de linguagem



Videoconferência



Modelagem e simulação

Superescalar e pipeline



Superescalar

Nesta organização, as instruções comuns (aritméticas de inteiros e pontos flutuantes), a carga do valor da memória em um registrador, o armazenamento do valor do registrador na memória e os desvios condicionais poderão ser iniciados e executados de forma independente.



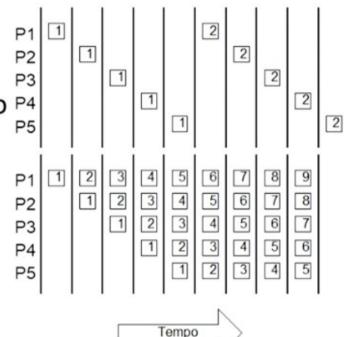
Pipeline



Nesta técnica, o caminho executável de uma instrução é dividido em estágios discretos, permitindo ao processador efetuar várias instruções simultaneamente, contanto que apenas uma delas ocupe cada estágio durante um ciclo de relógio.

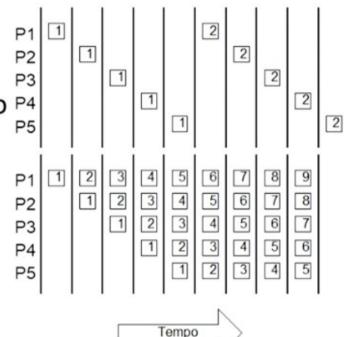
Pipeline

- P1 Unidade de busca de instruções
- P2 Unidade de decodificação de instrução P4
- P3 Unidade de busca de operando
- P4 Unidade de execução de instrução
- P5 Unidade de escrita



Pipeline

- P1 Unidade de busca de instruções
- P2 Unidade de decodificação de instrução P4
- P3 Unidade de busca de operando
- P4 Unidade de execução de instrução
- P5 Unidade de escrita





Arquitetura básica

> Como pensamos o processamento de um computador?

Atividade de pesquisa dicionário

- Processamento Paralelo*
- Taxonomia de Flynn
- Pipelines profundos
- Superpipelines*
- Superescalares
- Coerência de cache
- Arquitetura de acesso à memória
- Computadores multicores
- Aplicações multithead nativas*
- Aplicações com múltiplos processos