

# ARQUITETURA E ORGANIZAÇÃO DE *computadores*

Por Kauã de Jesus



KAUÃ DE JESUS

Arquitetura e Organização de  
Computadores, resumido em um  
livro

*Saiba do zero, em ordem de aprendizado, sobre  
arquitetura e organização de computadores*

*Copyright © 2023 by Kauā de Jesus*

*Kauā de Jesus asserts the moral right to be identified as the author of  
this work.*

*Designations used by companies to distinguish their products are often  
claimed as trademarks. All brand names and product names used in  
this book and on its cover are trade names, service marks, trademarks  
and registered trademarks of their respective owners. The publishers  
and the book are not associated with any product or vendor mentioned  
in this book. None of the companies referenced within the book have  
endorsed the book.*

*First edition*

*This book was professionally typeset on Reedsy.*

*Find out more at [reedsy.com](https://reedsy.com)*

# Contents

## I Introdução à arquitetura e organização de computadores

|  |   |
|--|---|
| 1 Conceitos básicos                                  | 3 |
| 2 História e evolução da arquitetura de computadores | 5 |
| 3 Visão geral dos componentes principais             | 8 |

## II Arquitetura de Von Neumann

|  |    |
|--|----|
| 4 Princípios básicos da arquitetura de Von Neumann | 13 |
| 5 Modelo de processamento                          | 15 |
| 6 Barramento do sistema                            | 17 |
| 7 Memória  | 19 |
| 8 Unidade de controle                              | 20 |

## III Organização da CPU

|                                      |    |
|--------------------------------------|----|
| 9 Estrutura e organização da CPU     | 23 |
| 10 Unidade lógica e aritmética (ULA) | 25 |
| 11 Registradores                     | 27 |
| 12 Pipeline de instruções            | 29 |
| 13 Microcódigo                       | 31 |
| 14 Cache de memória                  | 32 |

|      |  |    |
|------|--|----|
| IV   | Memória principal                              |    |
| 15   | Tipos de memória                               | 35 |
| 16   | Organização e hierarquia da memória            | 37 |
| 17   | Gerenciamento de memória                       | 39 |
| 18   | Acesso à memória                               | 41 |
| V    | Barramentos do sistema                         |    |
| 19   | Tipos de barramentos do sistema                | 45 |
| 20   | Barramento de endereço                         | 47 |
| 21   | Barramento de dados                            | 49 |
| 22   | Barramento de controle                         | 51 |
| VI   | Dispositivos de E/S (entrada e saída)          |    |
| 23   | Tipos de dispositivos de entrada e saída       | 55 |
| 24   | Comunicação entre a CPU e dispositivos de E/S  | 57 |
| 25   | Controle de E/S                                | 59 |
| 26   | Técnicas de download de dados                  | 61 |
| VII  | Interrupções e incorporadas                    |    |
| 27   | Conceitos básicos de controles e transferidos  | 65 |
| 28   | Como são usadas na arquitetura de computadores | 67 |
| 29   | Como o sistema operacional lida com elas       | 69 |
| VIII | Arquiteturas de processadores                  |    |
| 30   | RISC   | 73 |
| 31   | CISC   | 75 |

|                              |  |     |
|------------------------------|--|-----|
| 32                           | VLIW   | 76  |
| 33                           | EPIC   | 78  |
| IX Sistemas multiprocessados |  |     |
| 34                           | Conceitos básicos de sistemas multiprocessados               | 83  |
| 35                           | Tipos de sistemas multiprocessados                           | 85  |
| 36                           | Como eles funcionam  | 87  |
| 37                           | Como o sistema operacional gerencia o acesso à<br>CPU e à... | 89  |
| X Redes de computadores      |  |     |
| 38                           | Conceitos básicos de redes de computadores                   | 93  |
| 39                           | Topologias de rede   | 95  |
| 40                           | Protocolos de comunicação                                    | 97  |
| 41                           | Arquiteturas de rede   | 99  |
| <i>Also by Kauā de Jesus</i> |  | 101 |



# I

## Introdução à arquitetura e organização de computadores

*conceitos básicos, história e evolução da arquitetura de  
computadores, e uma visão geral dos componentes principais*



## Conceitos básicos

**A**rquitetura e organização de computadores é um campo de estudo que se preocupa com o projeto e implementação de sistemas computacionais. Esses sistemas são compostos por diversos componentes, como processadores, memória, dispositivos de entrada e saída, barramentos de sistema e redes de comunicação. A arquitetura e organização de computadores tem como objetivo garantir o funcionamento eficiente e confiável desses sistemas. Sabendo disso, aqui está a função básica de cada um desses componentes:

- Processadores: também conhecidos como CPUs (Central Processing Units), são o “cérebro” do computador, responsáveis por processar informações e executar programas. Eles são compostos por unidades lógicas e aritméticas, registradores e outras unidades que permitem que eles executem operações matemáticas, lógicas e de controle;
- Memória: é o componente que armazena dados e programas que estão sendo usados pelo processador. Existem diferentes tipos de memória, como RAM (Random Access Memory) e ROM (Read-Only Memory), cada uma com suas características específicas;

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

- Dispositivos de entrada e saída: são os componentes que permitem que o computador se comunique com o mundo exterior. Exemplos de dispositivos de entrada incluem teclados, mouses, scanners e microfones, enquanto exemplos de dispositivos de saída incluem monitores, impressoras e alto-falantes;
- Barramentos de sistema: são os canais que permitem que os diferentes componentes do computador se comuniquem entre si. Eles permitem que a CPU se comunique com a memória, dispositivos de entrada e saída e outros componentes do sistema;
- Redes de comunicação: permitem que diferentes computadores se comuniquem entre si, seja através de uma rede local (LAN) ou pela internet. As redes de comunicação são compostas por diferentes componentes, como roteadores, switches e cabos de rede;

Já temos uma boa base sabendo esses conceitos iniciais.

## História e evolução da arquitetura de computadores

**E**xiste um objeto que ainda é usado para brincadeiras de crianças e que foi considerado a primeira calculadora do mundo, conhecido como ‘ábaco’, criado aproximadamente em 4000 ac e era usado para cálculos simples e complexos . A partir disso, vieram algumas evoluções.

Pensando em computadores um pouco mais avançados, a história da arquitetura de computadores começou na década de 1930, com o desenvolvimento dos primeiros computadores mecânicos, como a máquina de calcular de Charles Babbage e Ada Lovelace.

### 1º Geração de computadores (1940 - 1956):

A primeira geração de computadores foi marcada pela utilização de válvulas eletrônicas como componentes principais. Esses computadores foram construídos na década de 1940 - durante a Segunda Guerra Mundial - e foram utilizados principalmente para fins militares e científicos. Eles eram grandes, caros e consumiam muita energia. Os programas eram criados por meio de linguagem de máquina e os dados eram inseridos por meio de cartões perfurados. Exemplos de

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

computadores da primeira geração incluem o ENIAC e o UNIVAC.

### 2° Geração de computadores (1956 - 1963):

A segunda geração de computadores teve início na década de 1950 e durou até meados dos anos 60. Nessa época, os transistores mudaram a substituir as válvulas usadas na primeira geração, tornando os computadores mais rápidos, menores e mais autônomos. Foi durante essa época que viveram os primeiros sistemas de processamento de lotes e sistemas operacionais, permitindo que os computadores executassem tarefas em um fluxo contínuo. A programação em linguagem Assembly também se tornou mais comum, permitindo que os programadores escrevam códigos mais eficientes e legíveis. Alguns dos computadores mais famosos da segunda geração incluem o IBM 1401 e o UNIVAC 1108.

### 3° Geração de computadores (1964 - 1971):

A terceira geração de computadores teve início em 1964 e foi marcada pela utilização de circuitos integrados (chips) em sua construção, o que permitia a criação de computadores mais potentes e com menor consumo de energia. Além disso, os computadores da terceira geração contavam com a utilização de sistemas operacionais, o que possibilitava o compartilhamento de recursos e o desenvolvimento de programas mais complexos. O IBM System/360, lançado em 1964, foi um dos principais computadores da terceira geração e se tornou um modelo padrão para a indústria de computadores. O advento dos microprocessadores, em meados da década de 1970, também marcou o fim da terceira geração e o início da quarta geração de computadores.

### 4° Geração de computadores (1971 - 1989):

A quarta geração de computadores começou em meados da década de 1970 e foi até o final dos anos 80. Essa geração de computadores foi

## HISTÓRIA E EVOLUÇÃO DA ARQUITETURA DE COMPUTADORES

identificada pelos microprocessadores, que integravam em um único chip todo o processamento central do computador. Isso permitiu a criação de computadores menores, mais baratos, mais eficientes e mais poderosos do que as gerações anteriores. Além disso, vivemos novas tecnologias, como as interfaces gráficas de usuário e as redes de computadores, que tornaram possível o uso dos computadores mais acessíveis e intuitivos para o público em geral.

### 5° Geração de computadores (1989 - dias atuais):

A quinta geração de computadores é caracterizada pelo desenvolvimento de tecnologias avançadas de inteligência artificial e linguagens de programação orientada a objetos. Essa geração começou no final dos anos 80 e início dos anos 90 e continua até hoje, com o desenvolvimento de novas tecnologias como a computação em nuvem e a internet das coisas. Além disso, houve um grande avanço na miniaturização dos componentes eletrônicos, o que possibilitou a criação de dispositivos cada vez menores e mais poderosos, como smartphones e tablets.

# 3

## Visão geral dos componentes principais

**A** arquitetura e organização de computadores envolve vários componentes principais, incluindo membrana (CPU), memória, dispositivos de entrada e saída, barramentos de sistema e redes de comunicação.

O processador (CPU) é o cérebro do computador, responsável por realizar operações lógicas e aritméticas e controlar o fluxo de dados. A memória é onde os dados e instruções são armazenados temporariamente para processamento pela CPU.

Os dispositivos de entrada e saída permitem que o usuário interaja com o computador, por meio de dispositivos como teclado, mouse, monitor, impressora, entre outros.

O barramento de sistema é responsável por conectar todos os componentes do computador e transferir dados entre eles. As redes de comunicação, por sua vez, permitem que vários computadores se comuniquem entre si, seja localmente ou pela internet.

A interação entre esses componentes é fundamental para o funcionamento do computador e para a execução de suas tarefas. A arquitetura e organização de computadores é uma área complexa e em constante evolução, com novas tecnologias e componentes sendo desenvolvidos

## VISÃO GERAL DOS COMPONENTES PRINCIPAIS

regularmente.



## II

# Arquitetura de Von Neumann

*princípios básicos da arquitetura von Neumann, modelo de processamento, barramento do sistema, memória e unidade de controle.*



# 4

## Princípios básicos da arquitetura de Von Neumann

**A**rquitetura von Neumann é uma arquitetura de computador que foi proposta por John von Neumann em 1945. Ela se tornou a base para a maioria dos computadores que conhecemos hoje.

Uma característica principal da Arquitetura von Neumann é a separação entre a unidade de processamento central (CPU), a memória e as entradas/saídas de dados. A CPU é responsável por processar as instruções que estão armazenadas na memória. A memória é utilizada para armazenar os dados e as instruções que serão processadas pela CPU. As entradas/saídas de dados são utilizadas para permitir a comunicação entre o computador e o mundo exterior, como teclado, mouse, monitor, impressora, entre outros.

Além disso, a Arquitetura von Neumann utiliza um barramento de dados para conectar uma CPU, uma memória e as entradas/saídas de dados. Isso permite que os dados sejam transferidos de forma eficiente entre esses componentes.

Outra característica importante da Arquitetura von Neumann é a utilização de um conjunto de instruções fixas, ou seja, uma CPU só é

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

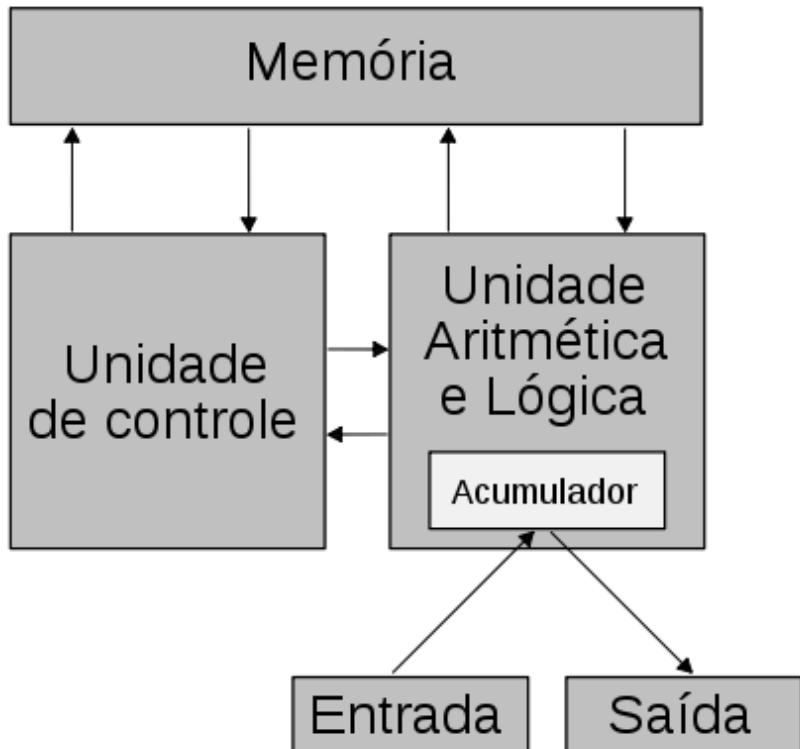
capaz de executar um conjunto limitado de instruções pré-definidas. Isso torna mais fácil para os programadores escreverem software, pois eles não precisam se preocupar em como as instruções serão executadas pelo hardware.

Esses princípios básicos da Arquitetura von Neumann ainda são utilizados em computadores modernos, embora tenham sido aprimorados com o tempo para torná-los mais eficientes e poderosos.

## Modelo de processamento

O modelo de processamento da arquitetura von Neumann é baseado em uma unidade central de processamento (CPU) que é responsável por executar as instruções armazenadas na memória principal do computador. O modelo é composto por quatro componentes principais: a unidade central de processamento (CPU), a memória principal, dispositivos de entrada/saída e barramento de sistema.

A CPU é responsável por buscar instruções da memória principal, decodificar as instruções e executá-las. A memória principal é utilizada para armazenar as instruções que a CPU deve executar, bem como os dados que serão processados. Os dispositivos de entrada/saída permitem que o computador se comunique com o mundo exterior, permitindo a entrada de dados e exibição de informações. O barramento de sistema é um conjunto de linhas de comunicação que permite que as informações sejam mantidas entre a CPU, a memória e os dispositivos de entrada/saída.



O modelo von Neumann também inclui a ideia de que as instruções e dados são armazenados na mesma memória, permitindo que a CPU acesse tanto as instruções quanto os dados necessários para executá-las. Essa arquitetura foi fundamental para o desenvolvimento dos computadores modernos e ainda é amplamente utilizada na maioria dos computadores pessoais e servidores.

# 6

## Barramento do sistema

O barramento do sistema é um componente fundamental da arquitetura de computadores, responsável por conectar e permitir a comunicação entre os diversos dispositivos do sistema, como CPU, memória, dispositivos de entrada e saída, dentre outros. Ele é responsável por transportar dados, endereços e sinais de controle, e é geralmente composto por três subsistemas principais: o barramento de dados, o barramento de endereços e o barramento de controle.

O barramento de dados é responsável por transportar os dados que são processados pela CPU, enquanto o barramento de endereços é responsável por transportar os endereços das posições de memória onde os dados serão lidos ou gravados. Já o barramento de controle é responsável por transportar os sinais de controle que determinam o momento em que os dados serão lidos ou gravados, além de outras operações de controle do sistema.

Os barramentos de sistema podem ter diferentes larguras de banda, que determinam a quantidade de dados que podem ser transportados em um determinado período de tempo. Quanto maior a largura de banda do barramento, mais rápido será o processamento de dados.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

Além disso, os barramentos de sistema também podem ter diferentes tipos de protocolos de comunicação, que determinam como os dados são transferidos e interpretados pelos dispositivos do sistema.

## Memória

**N**a arquitetura de von Neumann, a memória é uma das principais componentes do sistema. Ela é responsável por armazenar tanto os dados quanto as instruções que serão processadas pela CPU. A memória é organizada em endereços, e cada endereço contém um byte de informação.

A memória principal, também conhecida como memória RAM (Random Access Memory), é volátil, ou seja, ela perde seus dados quando a energia é difícil. A memória cache, por sua vez, é uma memória de acesso rápido que armazena temporariamente os dados mais usados pela CPU, observando assim o tempo de acesso à memória principal.

A arquitetura de von Neumann também prevê a utilização de memória secundária, como os discos rígidos (HD), que permitem o armazenamento de dados de forma permanente mesmo quando o computador está desligado. A transferência de dados entre a memória principal e a memória secundária é realizada por meio de controladores de disco, que controlam a leitura e gravação de dados nos discos.

## Unidade de controle

A unidade de controle é uma das principais unidades funcionais presentes na arquitetura de von Neumann. Ela é responsável por controlar a operação do processador e garantir que as instruções do programa sejam executadas corretamente.

A unidade de controle recebe as instruções da memória e as decodifica para determinar a operação que o processador deve realizar. Ela então coordena o fluxo de dados entre as unidades funcionais do processador, incluindo a unidade lógica e aritmética (ULA), a memória e os dispositivos de entrada/saída.

Para garantir que as instruções sejam executadas na ordem correta, a unidade de controle mantém um contador de programa que indica o endereço da próxima instrução a ser executada. Ela também pode incluir um registrador de estado que armazena informações sobre o estado atual do processador, como se houvesse uma operação de divisão por zero ou um estouro de memória.

Em resumo, a unidade de controle é responsável por coordenar as operações do processador e garantir que as instruções do programa sejam executadas corretamente e na ordem correta.

# III

## Organização da CPU

*estrutura e organização da CPU, incluindo a unidade lógica e aritmética (ULA), registradores, pipeline de instruções, microcódigo e cache de memória.*



## Estrutura e organização da CPU

**A**CPU (Central Processing Unit) é o componente principal de um computador responsável por executar as instruções dos programas. Ela é composta por três unidades principais: a unidade de controle, a unidade lógica aritmética e a unidade de registro.

A unidade de controle é responsável por interpretar as instruções do programa e controlar o fluxo de dados dentro do processador e entre o processador e a memória. Ela é composta por um contador de programa que armazena o endereço da próxima instrução a ser executada, um registrador de instrução que armazena a instrução atual e um decodificador de instrução que converte a instrução em sinais de controle para as outras unidades.

A unidade lógica aritmética (ULA) é responsável por executar operações lógicas e aritméticas, como adição, subtração, multiplicação e comparação. Ela é composta por um conjunto de circuitos eletrônicos que realizam essas operações e um registrador de resultado que armazena o resultado da operação.

A unidade de registro é responsável por armazenar dados temporários que estão sendo processados pela CPU. Ela é composta por um conjunto de registradores que armazenam dados em diferentes possibilidades de

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

processamento, como o registrador de entrada/saída, o registrador de endereço e o registrador de dados.

Além disso, a CPU também possui uma memória cache, que é uma memória de acesso rápido que armazena os dados mais usados pelo processador, mantendo o tempo de acesso à memória principal. A organização da CPU pode variar de acordo com a arquitetura do processador, mas essas unidades são comuns na grande maioria dos casos.

## Unidade lógica e aritmética (ULA)

A unidade lógica e aritmética (ULA) é um componente essencial da CPU, responsável por realizar as operações matemáticas e lógicas necessárias para o processamento de dados. A ULA é responsável por executar operações como adição, subtração, multiplicação, divisão, comparação e lógica booleana, entre outras.

A ULA recebe os dados a serem processados da memória ou dos registradores e, em seguida, executa uma operação especificada, produzindo um resultado que é armazenado de volta na memória ou em um registrador. Para realizar as operações aritméticas, a ULA usa circuitos digitais de somadores, enquanto para as operações lógicas, usa circuitos que implementam portas lógicas.

A ULA geralmente é organizada em dois blocos principais: a unidade lógica e a unidade aritmética. A unidade lógica é responsável por executar operações lógicas, como comparação e negação, enquanto a unidade aritmética é responsável por executar operações matemáticas, como adição e multiplicação. Alguns ULAs mais avançados podem integrar outras funcionalidades, como suporte a ponto flutuante.

A velocidade da ULA é um fator crítico no desempenho geral do processador. Por essa razão, muitos projetos de modernos têm múltiplas

ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

unidades de execução, incluindo múltiplas ULAs, para permitir a execução de várias instruções simultaneamente e aumentar o desempenho geral do processador.

## Registradores

**R**egistradores são elementos de memória de alta velocidade usados para armazenar dados temporários e instruções em uma unidade central de processamento (CPU) de um computador. Eles são usados para armazenar dados que precisam ser usados imediatamente pela CPU, como operandos de instruções, endereços de memória, resultados intermediários de operações aritméticas e lógicas e estados da CPU.

Os registradores são implementados em hardware na CPU e geralmente são acessados muito rapidamente, tornando-os ideais para armazenar dados temporários que são frequentemente acessados. Alguns exemplos comuns de registradores incluem:

- Registrador de acumulador: Usado para armazenar resultados intermediários de operações aritméticas e lógicas.
- Registradores de índice: Usados para indexar arrays e outras estruturas de dados na memória.
- Registradores de ponteiro: Usados para armazenar endereços de memória.
- Registradores de estado: Usados para armazenar informações sobre

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

o estado atual da CPU, como flags de condição e modo de operação.

Os registradores são projetados para trabalhar juntos de forma eficiente, permitindo que a CPU execute operações rapidamente. A unidade de controle da CPU usa os registradores para decodificar instruções e coordenar a execução de operações aritméticas e lógicas. À medida que os dados são processados, eles são movidos de um registrador para outro para permitir que a CPU execute várias operações em sequência.

Em resumo, os registradores são elementos críticos na organização da CPU de um computador, permitindo que a CPU processe dados rapidamente e execute operações complexas de forma eficiente.

## Pipeline de instruções

O pipeline de instruções é uma técnica utilizada em arquitetura de computadores para aumentar a eficiência do processamento de instruções pela CPU. Essa técnica consiste em dividir o processamento de uma instrução em várias etapas independentes e executá-las em paralelo.

O pipeline de instruções é composto por várias etapas que incluem a busca de instruções, a decodificação das instruções, a busca de operandos, a execução das instruções e o processamento dos resultados. Cada uma dessas etapas é executada por um circuito específico na CPU, chamado de estágio do pipeline.

Quando uma instrução é processada pelo pipeline, a próxima instrução é buscada e decodificada enquanto a instrução anterior está sendo executada. Dessa forma, várias instruções podem ser processadas em paralelo, aumentando a eficiência e eficiência o tempo de processamento.

No entanto, é importante lembrar que nem todas as instruções podem ser executadas em paralelo, e que a eficiência do pipeline pode ser experimentada por vários fatores, como dependências entre as instruções, conflitos de dados e instruções de salto condicional.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

Além disso, o pipeline de instruções pode ser implementado de diferentes modos, como o pipeline superescalar, que permite a execução de várias instruções em paralelo, ou o pipeline superpipelined, que divide cada estágio do pipeline em subestágios para aumentar ainda mais a eficiência.

## Microcódigo

**M**icrocode é uma sequência de instruções elementares armazenadas na memória de controle da CPU de um computador. Essas instruções são usadas para executar como instruções da linguagem de máquina de alto nível que a CPU recebe. O microcódigo é usado para interpretar as instruções da linguagem de máquina e para controlar o fluxo de dados e instruções pela CPU.

O microcódigo é especialmente útil para a implementação de instruções complexas e difíceis de executar diretamente no hardware da CPU. Ao utilizar o microcódigo, os treinos podem executar uma ampla gama de instruções, mesmo que o hardware físico não tenha a capacidade de executá-las.

Além disso, o microcódigo é usado para implementar funcionalidades adicionais, como controle de manutenção, monitoramento de desempenho e depuração de código. Em geral, o microcódigo é escrito por especialistas em microarquitetura e é considerado uma parte fundamental do projeto da CPU.

## Cache de memória

**C**ache de memória é uma pequena porção de memória de alta velocidade, que é usada para armazenar temporariamente os dados mais usados pelo processador, mantendo assim o tempo necessário para acessar a memória principal.

O cache é organizado em vários níveis (L1, L2, L3, etc.), sendo que o nível mais próximo do processador (L1) é o mais rápido, mas também o menor, enquanto que os níveis mais afastados são maiores, mas mais lentos.

Quando o processador precisa acessar um endereço de memória, ele primeiro verifica se esse endereço está no cache. Se estiver, o dado é recuperado diretamente do cache, sem a necessidade de acessar a memória principal. Se não estiver no cache, o dado é recuperado da memória principal e copiado para o cache, para que possa ser acessado mais rapidamente em operações futuras.

O cache é gerenciado pelo hardware do processador, de forma transparente para o sistema operacional e para os programas em execução. Para maximizar a eficiência do cache, é importante que o software seja escrito de forma a maximizar o uso de dados que estão presentes no cache.

# IV

## Memória principal

*tipos de memória, organização e hierarquia da memória,  
gerenciamento de memória e acesso à memória.*



## Tipos de memória

A memória principal é um dos componentes essenciais da arquitetura e organização de computadores. É responsável por armazenar temporariamente os dados e instruções que o processador precisa acessar para executar tarefas. A seguir, estão alguns dos tipos de memória principal comumente encontrados:

1. Memória de acesso aleatório (RAM): é uma memória volátil que permite a leitura e escrita de dados. É a forma mais comum de memória usada em computadores e é geralmente classificada pelo tipo de tecnologia utilizada, como DRAM (Dynamic RAM) ou SRAM (Static RAM).
2. Memória somente de leitura (ROM): é uma memória não volátil que contém dados que não podem ser modificados. É usado para armazenar o firmware do sistema, como o BIOS.
3. Memória flash: é uma forma de memória não volátil que é regravável. É comumente encontrado em dispositivos de armazenamento externos, como unidades USB, cartões SD e unidades de estado sólido.
4. Cache: é uma forma especial de memória de acesso rápido que

armazena dados frequentemente acessados para acelerar o tempo de acesso à memória principal. Existem vários níveis de cache, cada um com diferentes recursos e velocidades de acesso.

5. Memória virtual: é uma técnica usada pelo sistema operacional para simular a existência de mais memória do que realmente está disponível. Isso é feito movendo dados da memória principal para o disco rígido quando a memória principal está cheia.

Em geral, o tipo de memória usado em um computador depende de fatores como velocidade de acesso, custo e volatilidade. O uso de diferentes tipos de memória é uma parte importante do projeto de sistemas de computadores eficientes e controlados.

## Organização e hierarquia da memória

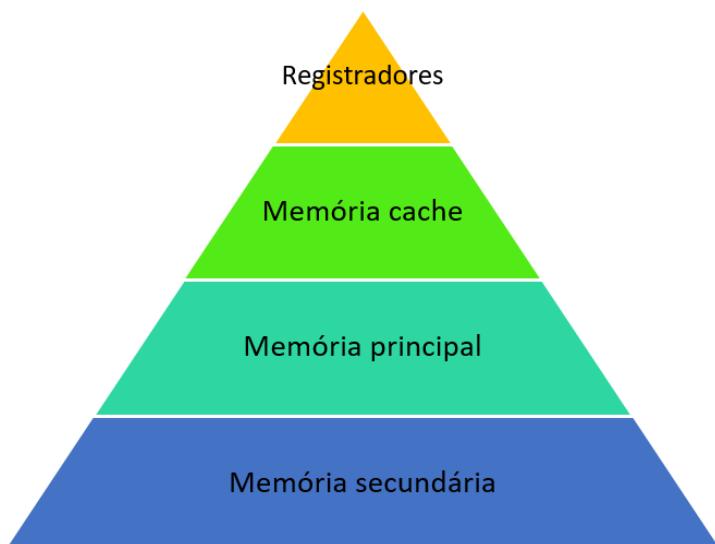
A organização e hierarquia da memória é uma estrutura que define como os dados são armazenados e acessados no sistema de computação. A hierarquia é composta por diferentes níveis de memória, que variam em termos de capacidade, velocidade e custo.

O primeiro nível da autoridade é a memória cache, que é a memória mais rápida e cara do sistema. Um cache é usado para armazenar dados e instruções que são frequentemente usados pelo processador. Ela é dividida em níveis, como L1, L2 e L3, e quanto mais próximo do processador, mais rápido é o acesso à cache.

O segundo nível da autoridade é a memória principal, também conhecida como RAM. A RAM é mais lenta que o cache, mas tem uma capacidade maior. É usado para armazenar temporariamente dados e instruções que estão sendo usados pelo processador no momento. A RAM pode ser volátil ou não volátil, ou seja, pode ou não reter dados quando a energia é difícil.

O terceiro nível da hierarquia é o armazenamento secundário, que é usado para armazenar dados permanentemente, como discos rígidos, SSDs e pendrives. Esses dispositivos têm uma capacidade muito maior que a RAM, mas são mais lentos.

A hierarquia de memória é importante porque permite que os dados sejam acessados rapidamente quando necessário, enquanto mantém uma capacidade de armazenamento razoável e um custo viável. O sistema de gerenciamento de memória é responsável por mover os dados entre os diferentes níveis da hierarquia, a fim de garantir o melhor desempenho possível.



## Gerenciamento de memória

O gerenciamento de memória é uma função importante do sistema operacional que gerencia a alocação de espaço de memória para os programas em execução. Existem duas técnicas principais de gerenciamento de memória: alocação de memória estática e dinâmica.

Na alocação de memória estática, o espaço de memória é alocado permanentemente para um programa no momento da compilação ou do carregamento. Isso significa que o espaço de memória atribuído ao programa permanece o mesmo durante toda a execução.

Na alocação de memória dinâmica, o espaço de memória é alocado e desalocado conforme necessário durante a execução do programa. Isso permite que vários programas possam compartilhar a mesma memória física e aumentar a eficiência do sistema, pois não há desperdício de memória.

O gerenciamento de memória pode ser implementado de várias maneiras, incluindo:

- Memória dividida em páginas: a memória é dividida em páginas de tamanho fixo e o sistema operacional mantém uma tabela de

páginas que mapeia o espaço de endereço virtual do programa para o espaço de endereço físico da memória.

- Memória dividida em segmentos: a memória é dividida em segmentos de tamanho variável, onde cada segmento contém um tipo específico de dados, como código executável, dados do programa e pilha.
- Memória virtual: cada programa é atribuído a um espaço de endereço virtual que é mapeado para um espaço de endereço físico pela unidade de gerenciamento de memória. Isso permite que o sistema operacional aloque espaço de memória de forma mais eficiente, pois o espaço de endereço virtual pode ser maior do que o espaço de endereço físico disponível na memória.

Além disso, existem vários algoritmos de substituição de página que são usados para decidir qual página deve ser removida da memória quando a memória está cheia e uma nova página precisa ser carregada. Alguns exemplos desses algoritmos incluem o algoritmo FIFO (primeiro a entrar, primeiro a sair), LRU (menos recentemente usado) e o algoritmo de substituição de página ótimo.

## Acesso à memória

O acesso à memória principal em um computador é feito por meio de um endereço de memória. O endereço de memória é um valor numérico que especifica a localização da célula de memória que se deseja acessar. O processo de leitura e gravação na memória é feito por meio de instruções que são executadas pela CPU.

Para garantir a eficiência do acesso à memória, a memória é organizada em blocos de tamanho fixo chamados de palavras de memória. A largura da palavra de memória é determinada pela arquitetura do processador e geralmente varia entre 8 e 64 bits. A largura da palavra de memória é importante porque determina o número de bits que podem ser lidos ou gravados de uma só vez.

Existem duas formas principais de acesso à memória: acesso direto e acesso indireto. No acesso direto, a CPU acessa a memória diretamente usando o endereço de memória. Já no acesso indireto, a CPU usa um registrador especial chamado de registrador de endereço de memória para armazenar o endereço de memória.

O acesso à memória também pode ser feito por meio de barramentos de dados e endereço. O barramento de dados é usado para transferir dados entre a CPU e a memória, enquanto o barramento de endereço é

usado para especificar o endereço de memória a ser acessado.

O gerenciamento de memória também é importante para garantir que cada programa tenha acesso à quantidade necessária de memória. Isso é feito por meio de técnicas de alocação de memória, como o particionamento de memória e a paginação de memória. O particionamento de memória divide a memória em partições fixas, enquanto a paginação de memória divide a memória em páginas de tamanho fixo. Ambas as técnicas são usadas para otimizar o uso da memória disponível e minimizar os problemas de fragmentação da memória.

# V

## Barramentos do sistema

*tipos de barramentos do sistema, incluindo barramento de endereço, barramento de dados e barramento de controle.*



## Tipos de barramentos do sistema

**O**s barramentos do sistema são componentes essenciais da arquitetura de um computador, responsáveis por permitir a comunicação entre diferentes componentes do sistema. Existem diferentes tipos de barramentos, cada um com uma concepção específica.

- Barramento de dados: é o barramento responsável por transmitir os dados entre a CPU e a memória. Ele tem largura determinada pela arquitetura do sistema, e pode variar entre 8, 16, 32 ou mais bits.
- Barramento de endereço: é o barramento responsável por indicar o endereço da memória ou dispositivo que deseja acessar. Ele tem largura determinada pela quantidade de endereços que podem ser gerados pela CPU.
- Barramento de controle: é o barramento responsável por transmitir sinais de controle entre a CPU e outros dispositivos, como a memória, dispositivos de entrada/saída, etc. Ele carrega sinais que controlam o fluxo de dados e de endereços, sinais de interrupção, sinais de clock, entre outros.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

- Barramento de expansão: é o barramento usado para conectar dispositivos de expansão, como placas de rede, placas de vídeo, entre outros. Ele permite que esses dispositivos sejam adicionados ao sistema, aumentando seus recursos.

Cada barramento tem uma amplitude e frequência de operação específica, que dependem da arquitetura do sistema. Além disso, a utilização de barramentos de alta velocidade e largura de banda é essencial para garantir o desempenho de sistemas modernos, especialmente em aplicações que comportam grande volume de dados e processamento.

## Barramento de endereço

O barramento de endereço é um componente do sistema de computador responsável por transmitir os sinais de endereço de memória e dispositivos de entrada/saída para o controlador de memória e E/S. Ele permite que o processador possa acessar a memória principal e dispositivos de entrada/saída para buscar e armazenar informações.

O barramento de endereço é unidirecional, ou seja, somente transmite os sinais de endereço do processador para a memória ou E/S. Ele tem um tamanho que determina quantos bits de endereço podem ser transmitidos ao mesmo tempo, o que influencia diretamente na quantidade máxima de memória que o sistema pode endereçar. Por exemplo, um barramento de endereço de 16 bits pode endereçar no máximo 64KB de memória.

A velocidade de transmissão do barramento de endereço também é importante, pois afeta diretamente o desempenho do sistema. Quanto maior a velocidade, mais rápido o processador pode acessar a memória e dispositivos de E/S.

O barramento de endereço é um dos três tipos principais de barramentos do sistema, juntamente com o barramento de dados e o barramento

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

de controle. Juntos, eles garantem uma comunicação eficiente entre os componentes do sistema de computador.

## Barramento de dados

O barramento de dados é um componente da arquitetura de computadores responsável pelo transporte de informações entre os componentes do sistema, como a CPU, memória e dispositivos de entrada/saída. Especificamente, ele é responsável pelo transporte dos dados que estão sendo transferidos entre esses componentes.

O barramento de dados tem uma largura definida, que indica quantos bits de dados podem ser transferidos simultaneamente. Por exemplo, um barramento de dados de 32 bits pode transferir 32 bits de dados por vez. A largura do barramento de dados afeta a velocidade de transferência de dados, pois quanto maior a largura, mais dados podem ser transferidos de uma só vez.

Além disso, o barramento de dados pode ser unidirecional ou bidirecional. Em um barramento de dados unidirecional, os dados são transferidos apenas em uma direção, enquanto em um barramento de dados bidirecional, os dados podem ser transferidos em ambas as direções.

O barramento de dados é um dos três principais tipos de barramentos do sistema, juntamente com o barramento de endereço e o barramento

ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

de controle. Juntos, esses barramentos coordenam a comunicação entre a CPU, a memória e os dispositivos de entrada/saída em um sistema de computador.

## Barramento de controle

O barramento de controle é responsável por transmitir sinais de controle entre os componentes do computador, como a CPU, a memória e os dispositivos de entrada/saída. Esse barramento é responsável por sincronizar o funcionamento dos componentes do sistema, permitindo que eles se comuniquem e coordeneem suas atividades.

Os sinais de controle incluem comandos como “ler”, “escrever”, “iniciar” e “parar”. Através do barramento de controle, a CPU pode enviar esses comandos para a memória ou para os dispositivos de entrada/saída, instruindo-os a realizar determinadas operações.

O barramento de controle é geralmente dividido em sub-barramentos, cada um dos quais transmite um tipo específico de sinal de controle. Esses sub-barramentos incluem o barramento de interrupção, que é usado para lidar com interrupções no sistema, o barramento de clock, que transmite sinais de sincronização para todos os componentes do sistema, e o barramento de status, que transmite informações de status de volta à CPU.



# VI

## Dispositivos de E/S (entrada e saída)

*tipos de dispositivos de entrada e saída, comunicação entre a CPU e dispositivos de E/S, controle de E/S e técnicas de download de dados.*



## Tipos de dispositivos de entrada e saída

**O**s dispositivos de entrada e saída (E/S) são componentes que permitem que o computador se comunique com o mundo exterior. Eles são responsáveis por permitir a entrada de informações no computador e a saída de informações para o usuário ou outros dispositivos externos.

Alguns exemplos de dispositivos de entrada incluem:

- Teclado: permite a entrada de dados através de teclas;
- Mouse: permite a entrada de comandos e movimentos do cursor na tela;
- Scanner: permite a digitalização de documentos e imagens;
- Microfone: permite a entrada de som e voz;
- Câmera: permite a entrada de imagens e vídeo;
- Leitor de código de barras: permite a leitura de informações em códigos de barras.

Alguns exemplos de dispositivos de saída incluem:

- Monitor: exibe imagens e informações na tela;

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

- Impressora: imprime documentos e imagens;
- Caixas de som: reproduzem áudio;
- Projetor: exibe imagens e informações em uma superfície maior;
- Alto-falantes: reproduzem áudio em alta qualidade.

Os dispositivos de E/S são conectados ao computador por meio de portas específicas, como USB, HDMI, VGA, entre outras. Eles podem ser internos, como é o caso de placas de som e rede, ou externos, como é o caso de mouses e teclados conectados por cabo ou sem fio.

## Comunicação entre a CPU e dispositivos de E/S

**A** comunicação entre a CPU e os dispositivos de entrada e saída (E/S) é feita através de controladores de E/S e barramentos de E/S. Os controladores de E/S são responsáveis por gerenciar a comunicação entre a CPU e os dispositivos, realizando as operações de entrada e saída, incluindo a conversão de dados entre formatos internos do dispositivo e do sistema.

Os barramentos de E/S são os meios físicos pelos quais a CPU se comunica com os dispositivos de E/S. Existem diferentes tipos de barramentos de E/S, dependendo da finalidade e do tipo de dispositivo. Por exemplo, barramentos USB, HDMI, SATA, Ethernet, entre outros.

Ao receber uma solicitação de E/S, a CPU envia um sinal de interrupção para o controlador de E/S. Esse sinal interrompe a operação da CPU e indica que há uma solicitação de E/S que precisa ser atendida. O controlador de E/S, por sua vez, coordena a operação do dispositivo de E/S, permitindo a transferência de dados entre o dispositivo e a CPU.

A comunicação entre a CPU e os dispositivos de E/S pode ocorrer de forma síncrona ou assíncrona. Na comunicação síncrona, a CPU espera até que a operação de E/S seja concluída antes de continuar a executar

ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

outras instruções. Na comunicação assíncrona, a CPU continua a executar outras instruções enquanto aguarda a conclusão da operação de E/S.

## Controle de E/S

O controle de E/S (Entrada e Saída) é uma das funções mais importantes do sistema de computação, pois permite que os dispositivos externos se comuniquem com a CPU. É um processo complexo que envolve a coordenação entre a CPU, os dispositivos de E/S e o sistema de memória.

Existem duas formas principais de realizar a E/S em um computador: programada e por interrupção. Na E/S programada, a CPU é responsável por controlar todas as operações de E/S, incluindo a transferência de dados entre a memória principal e o dispositivo de E/S. Já na E/S por interrupção, o dispositivo de E/S envia um sinal de interrupção para a CPU para avisá-la que uma operação de E/S foi concluída. A CPU então interrompe a execução do programa atual e começa a executar uma rotina de serviço de interrupção para lidar com a operação de E/S.

O controle de E/S também envolve o gerenciamento de dispositivos de E/S, como impressoras, teclados, mouses, monitores, entre outros. Cada dispositivo de E/S é identificado por um endereço específico e um conjunto de registradores de controle. A CPU usa esses endereços e registradores para enviar e receber dados dos dispositivos de E/S.

Além disso, o controle de E/S também inclui o gerenciamento de

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

buffers de entrada e saída, que são áreas de memória reservadas para armazenar temporariamente dados que estão sendo transferidos entre a CPU e os dispositivos de E/S. Isso ajuda a evitar gargalos e a aumentar a eficiência do sistema de E/S.

Em resumo, o controle de E/S é essencial para permitir a comunicação entre a CPU e os dispositivos de E/S, e envolve o gerenciamento de dispositivos, buffers e a coordenação entre os diferentes componentes do sistema de computação.

## Técnicas de download de dados

**A**s técnicas de download de dados são utilizadas para transferir informações de um servidor para um dispositivo local. Existem diferentes métodos de download de dados, cada um com suas vantagens e desvantagens. Alguns dos métodos mais comuns incluem:

1. Download direto: este é o método mais simples de download de dados, em que o arquivo é baixado diretamente do servidor para o dispositivo local. É amplamente utilizado para baixar arquivos de tamanho pequeno a médio, como documentos, imagens e vídeos curtos.
2. Transferência FTP: o FTP (File Transfer Protocol) é um protocolo de rede que permite a transferência de arquivos entre um servidor e um cliente. É frequentemente utilizado para transferir grandes quantidades de dados, como backups de banco de dados, arquivos de mídia de grande porte, entre outros.
3. Transferência HTTP: a transferência de dados por meio do protocolo HTTP (Hypertext Transfer Protocol) é uma técnica de download de dados muito utilizada para arquivos de tamanho

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

pequeno a médio, como imagens, documentos e arquivos de áudio e vídeo de curta duração.

4. Streaming de mídia: streaming de mídia é uma técnica de download de dados em que os dados são entregues ao dispositivo em tempo real, permitindo que o usuário assista a um vídeo ou ouça uma música enquanto o arquivo é baixado.
5. Peer-to-peer (P2P): o compartilhamento de arquivos P2P é uma técnica em que os dados são transferidos diretamente entre dois ou mais dispositivos. Este método é frequentemente utilizado para transferir arquivos grandes, como arquivos de filmes e jogos.
6. Download em lote: esta técnica de download de dados permite que os usuários baixem vários arquivos simultaneamente. É especialmente útil para arquivos de tamanho pequeno a médio, como imagens e documentos.

Independentemente do método utilizado, é importante ter em mente as políticas de uso aceitável e direitos autorais ao baixar dados da internet.

## VII

### Interrupções e incorporadas

*conceitos básicos de controles e transferidos, como elas são usadas na arquitetura de computadores e como o sistema operacional lida com elas.*



## Conceitos básicos de controles e transferidos

**I**nterrupção e dispositivos incorporados são conceitos importantes na área de sistemas embarcados e microcontroladores.

A interrupção é uma técnica utilizada em computação para suspender temporariamente a execução do programa principal e tratar eventos que ocorrem em paralelo, como uma requisição de entrada/saída ou uma interação com o usuário. Quando uma interrupção é gerada, o processador interrompe sua tarefa atual e passa a executar uma rotina de serviço (chamada de rotina de tratamento de interrupção), que lida com a solicitação de interrupção. Depois que a rotina de serviço é concluída, o processador retorna à tarefa anterior.

Dispositivos incorporados são dispositivos eletrônicos que possuem uma ou mais funções dedicadas, como sensores, atuadores, controles, displays, entre outros. Eles são projetados para executar tarefas específicas e geralmente estão embutidos em um sistema maior, como em um microcontrolador. A comunicação entre a CPU e os dispositivos incorporados é geralmente feita através de portas de entrada/saída.

No controle de dispositivos incorporados, a interrupção é usada para notificar o processador quando um dispositivo requer atenção

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

ou quando um evento ocorre, como a conclusão de uma operação de entrada/saída. Em muitos casos, a interrupção é usada para despertar o processador do modo de baixo consumo, economizando energia. A comunicação entre a CPU e os dispositivos incorporados é geralmente realizada através de portas de entrada/saída e protocolos de comunicação dedicados.

## Como são usadas na arquitetura de computadores

**A**s interrupções e as rotinas incorporadas são usadas na arquitetura de computadores para permitir que o processador execute tarefas diversas, de forma eficiente e coordenada.

As interrupções são sinais emitidos por dispositivos externos ao processador, indicando a necessidade de atenção ou a ocorrência de um evento específico. O processador interrompe sua tarefa atual, salva o estado atual do programa e começa a executar uma rotina de serviço, que pode estar localizada na ROM ou na RAM. Essa rotina é responsável por tratar a interrupção e pode executar uma série de operações específicas, como receber dados de um dispositivo, enviar dados para uma impressora, atualizar a tela, entre outras.

As rotinas incorporadas, por outro lado, são trechos de código pré-programados no processador, que realizam operações específicas de baixo nível. Elas são utilizadas para operações rotineiras do sistema, como realizar operações matemáticas, mover dados entre registradores, ou ainda, controlar a execução do programa.

Ambas as técnicas permitem que o processador execute diversas tarefas ao mesmo tempo, de forma coordenada e eficiente, aumentando

ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM  
LIVRO

a capacidade e a velocidade de processamento do sistema.

## Como o sistema operacional lida com elas

**A**s interrupções e exceções são mecanismos fundamentais que permitem que o sistema operacional interaja com os dispositivos de E/S e gerencie os recursos do sistema de maneira eficiente. Quando uma interrupção ou exceção ocorre, o sistema operacional é notificado e executa um tratador de interrupção ou exceção.

O tratador de interrupção é uma rotina especial que é invocada automaticamente pelo sistema operacional quando uma interrupção é detectada. Essa rotina lida com a interrupção e determina a ação apropriada a ser tomada. Por exemplo, quando um dispositivo de E/S envia um sinal de interrupção para a CPU, o tratador de interrupção correspondente é acionado e pode ler os dados do dispositivo ou enviar dados para ele, dependendo do tipo de operação que foi realizada.

As exceções são erros ou situações anormais que ocorrem durante a execução de um programa. Algumas exceções comuns incluem falhas de segmentação (quando um programa tenta acessar uma parte da memória que não lhe pertence), divisão por zero e erros de ponto flutuante. Quando uma exceção ocorre, o sistema operacional interrompe a execução do programa e chama o tratador de exceção correspondente.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

Esse tratador pode tentar corrigir o erro ou finalizar o programa se o erro for fatal.

Em resumo, as interrupções e exceções são essenciais para o funcionamento do sistema operacional e permitem que ele gerencie os recursos do sistema de maneira eficiente e lide com situações anormais que podem ocorrer durante a execução dos programas.

## VIII

# Arquiteturas de processadores

*diferentes tipos de arquiteturas de processadores, incluindo RISC, CISC, VLIW e EPIC.*



## RISC

**A**rquitetura RISC (Reduced Instruction Set Computer ou Computador com Conjunto de Instruções Reduzido) é uma abordagem de design de processadores que se concentra em reduzir a complexidade das instruções para aumentar o desempenho e a eficiência energética. Em vez de ter instruções complexas que realizam várias operações, a arquitetura RISC se concentra em instruções simples que realizam operações básicas, mas de forma mais rápida.

Algumas das características da arquitetura RISC incluem:

- Conjunto de instruções reduzido: a maioria das instruções executam uma única operação simples, em vez de instruções complexas que executam várias operações.
- Instruções de tamanho fixo: cada instrução é codificada em um tamanho fixo de bits, o que simplifica o processo de decodificação e execução das instruções.
- Acesso à memória apenas através de instruções específicas: isso simplifica o gerenciamento da memória e reduz a complexidade do processador.
- Uso extensivo de registradores: a arquitetura RISC tem um grande

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

número de registradores, permitindo que mais operações sejam realizadas sem acessar a memória.

- Pipelining: a arquitetura RISC é frequentemente usada em processadores que implementam a técnica de pipeline, que permite que várias instruções sejam executadas simultaneamente.

A arquitetura RISC é usada em uma variedade de processadores, incluindo ARM, PowerPC e MIPS. É frequentemente usada em dispositivos móveis e sistemas embarcados, bem como em supercomputadores e servidores de alta performance.

# 31

## CISC

**A** arquitetura CISC (Complex Instruction Set Computer) é uma abordagem de projeto de processadores que visa a redução do número de instruções executadas por um programa. Em contraste com a arquitetura RISC (Reduced Instruction Set Computer), a CISC tem como característica o uso de instruções mais complexas e completas, capazes de realizar múltiplas operações em um único ciclo de clock.

Diferentemente da arquitetura RISC, que tem como princípio a simplicidade das instruções e a execução em ciclos mais curtos, a CISC busca oferecer ao programador um conjunto mais extenso de instruções, com capacidade de processamento mais diversificada. Isso significa que o processador CISC é capaz de executar instruções mais complexas em menos ciclos de clock, o que, em tese, aumenta o desempenho do sistema.

Alguns dos exemplos de processadores baseados na arquitetura CISC são a família Intel x86, utilizada em computadores pessoais, e a arquitetura ARM, usada em dispositivos móveis e sistemas embarcados.

## VLIW

**A** arquitetura VLIW (Very Long Instruction Word) é um tipo de arquitetura de processador que tenta maximizar a utilização da unidade de processamento do processador, permitindo que múltiplas instruções sejam executadas em paralelo. A arquitetura VLIW é uma abordagem de projeto alternativa à arquitetura RISC (Reduced Instruction Set Computing) e à arquitetura CISC (Complex Instruction Set Computing).

Na arquitetura VLIW, o compilador é responsável por gerar instruções muito longas (VLIW) que contêm várias instruções independentes que podem ser executadas simultaneamente. Cada VLIW é dividido em “slots”, e cada slot contém uma instrução independente que pode ser executada simultaneamente com as outras instruções no VLIW. O processador VLIW é projetado para executar várias instruções simultaneamente em ciclos de clock únicos, sem a necessidade de uma unidade de execução complexa.

Uma das principais vantagens da arquitetura VLIW é a simplicidade do hardware do processador, uma vez que não precisa de um hardware complexo de pipeline, ou mesmo de uma unidade de execução complexa. Isso permite que a arquitetura VLIW seja facilmente escalável, e pode

ser usada em processadores de alta performance para aplicações como processamento de imagens, processamento de sinal digital, ou mesmo processamento de vídeo.

No entanto, uma das principais desvantagens da arquitetura VLIW é a dificuldade de criar instruções longas e independentes que possam ser executadas simultaneamente sem conflitos. Isso faz com que o desempenho da arquitetura VLIW seja altamente dependente do compilador, e muitas vezes os compiladores precisam ser altamente otimizados para aproveitar completamente o potencial da arquitetura VLIW. Além disso, a arquitetura VLIW geralmente consome mais energia do que a arquitetura RISC, o que pode ser uma preocupação para aplicações de baixo consumo de energia.

## EPIC

**A** arquitetura EPIC (Explicitly Parallel Instruction Computing) é uma arquitetura de processador desenvolvida pela Intel e HP (Hewlett-Packard). Ela se baseia em uma estratégia de processamento de instruções que visa a execução paralela de um grande número de instruções por ciclo de clock, utilizando recursos de hardware especializados para garantir a execução correta e eficiente das instruções.

Ao contrário de outras arquiteturas, como a CISC e a RISC, que utilizam uma abordagem de decodificação de instruções em tempo real, a arquitetura EPIC utiliza um compilador avançado para gerar instruções altamente paralelas e otimizadas. Isso permite que o processador execute várias instruções simultaneamente, sem a necessidade de decodificar instruções complexas em tempo real.

Os processadores EPIC são projetados com recursos de execução especiais, como múltiplas unidades de execução, unidades de controle de fluxo, buffers de reordenação de instruções e previsão de desvio avançada. Esses recursos permitem que as instruções sejam executadas em paralelo, mesmo quando há dependências entre elas.

Outra característica importante da arquitetura EPIC é a utilização

de bundles de instruções. Os bundles são grupos de até três instruções relacionadas que são executadas juntas em um único ciclo de clock. Cada bundle contém instruções independentes, o que permite que o processador maximize a utilização dos recursos de hardware disponíveis.

Em resumo, a arquitetura EPIC é uma abordagem de processamento altamente paralela e otimizada que utiliza um compilador avançado para gerar instruções altamente paralelas e otimizadas. Ela é projetada para executar várias instruções simultaneamente, sem a necessidade de decodificação de instruções complexas em tempo real, e utiliza recursos de hardware especializados para garantir a execução correta e eficiente das instruções.



# IX

## Sistemas multiprocessados

*conceitos básicos de sistemas multiprocessados, tipos de sistemas multiprocessados, como eles funcionam e como o sistema operacional gerencia o acesso à CPU e à memória.*



## Conceitos básicos de sistemas multiprocessados

Sistemas multiprocessados são aqueles que possuem mais de um processador (CPU) trabalhando juntos em um único sistema para melhorar o desempenho e aumentar a capacidade de processamento. Existem duas principais abordagens para sistemas multiprocessados: sistemas simétricos (SMP) e sistemas assimétricos (AMP).

Em um sistema multiprocessado simétrico, todos os processadores têm acesso igual à memória e aos dispositivos de entrada e saída. Cada processador é capaz de executar qualquer tarefa, e o sistema operacional gerencia a alocação de tarefas para os diferentes processadores. Os processadores em um sistema SMP são interconectados por meio de um barramento de sistema compartilhado, que é responsável por transferir dados entre os processadores, a memória e os dispositivos de entrada e saída.

Em um sistema multiprocessado assimétrico, existem um ou mais processadores que são mais poderosos do que os outros e são designados para executar tarefas específicas. Esses processadores são chamados de “mestres” e são responsáveis por gerenciar as tarefas em execução no

sistema. Os processadores em um sistema AMP geralmente têm sua própria memória dedicada e são interconectados por meio de uma rede de alta velocidade.

Os sistemas multiprocessados podem melhorar o desempenho em relação aos sistemas de processador único, permitindo que várias tarefas sejam executadas simultaneamente. Além disso, eles oferecem maior confiabilidade e disponibilidade, pois, em caso de falha de um processador, os outros processadores podem assumir a carga de trabalho. No entanto, eles também apresentam desafios de projeto, como a necessidade de gerenciar a comunicação entre os processadores e garantir a consistência dos dados compartilhados.

## Tipos de sistemas multiprocessados

**E**xistem basicamente dois tipos de sistemas multiprocessados:

1. Sistemas multiprocessados simétricos (SMP): nesses sistemas, todos os processadores têm acesso a toda a memória e aos dispositivos de entrada/saída. Eles são usados para executar programas paralelos que requerem alta capacidade de processamento e baixa latência de memória. Nesse tipo de sistema, a tarefa de escalonamento é responsabilidade do sistema operacional, que decide qual processador executará cada tarefa.
2. Sistemas multiprocessados assimétricos (AMP): nesses sistemas, cada processador é projetado para executar tarefas específicas. Eles são usados em aplicativos que exigem alta confiabilidade e disponibilidade, como sistemas de controle de tráfego aéreo e sistemas de defesa. Nesse tipo de sistema, as tarefas são atribuídas a processadores específicos e não há necessidade de escalonamento pelo sistema operacional.

Além disso, existem os sistemas de processamento distribuído, que

ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

são compostos por várias máquinas independentes que se comunicam por meio de uma rede e trabalham juntas para resolver um problema específico. Esses sistemas são usados em aplicativos de grande escala, como bancos de dados distribuídos e sistemas de processamento de imagem.

## Como eles funcionam

Sistemas multiprocessados são sistemas que possuem mais de um processador, podendo ser utilizados de forma simultânea para a execução de tarefas. Existem diversas arquiteturas e modelos de sistemas multiprocessados, mas de forma geral, eles funcionam de maneira coordenada e sincronizada, compartilhando recursos como memória, barramentos e dispositivos de entrada/saída.

Em sistemas multiprocessados simétricos (SMP), todos os processadores têm acesso igualitário a memória e dispositivos de entrada/saída, permitindo que qualquer processador execute qualquer tarefa. Já em sistemas multiprocessados assimétricos (ASMP), cada processador é destinado a tarefas específicas, podendo ser divididos em processadores de entrada/saída, processadores de cálculo e processadores de controle, por exemplo.

Para garantir que os processadores trabalhem de forma sincronizada e coordenada, é necessário um sistema de comunicação entre eles, geralmente por meio de barramentos e sinais de controle. Além disso, é necessário um sistema de gerenciamento de tarefas que distribua as tarefas entre os processadores e garanta que não ocorram conflitos no acesso aos recursos compartilhados.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

Os sistemas multiprocessados oferecem vantagens em relação a sistemas com apenas um processador, como maior desempenho e capacidade de processamento paralelo. No entanto, a implementação de sistemas multiprocessados pode ser mais complexa e demandar um maior investimento em hardware e software.

## Como o sistema operacional gerencia o acesso à CPU e à memória

**E**m sistemas multiprocessamento, o sistema operacional é responsável por gerenciar o acesso à CPU e à memória pelos processos em execução. Existem diversas técnicas de gerenciamento de processos e escalonamento de CPU, incluindo:

1. Round-robin: O escalonador do sistema operacional dá a cada processo um tempo de CPU fixo, chamado de quantum, antes de passar para o próximo processo.
2. Prioridade: Os processos são atribuídos uma prioridade e o escalonador dá preferência aos processos com prioridade mais alta.
3. Deadlines: Os processos têm uma data de conclusão definida e o escalonador gerencia a CPU para garantir que os processos sejam concluídos dentro do prazo.
4. Escalonamento de múltiplas filas: Os processos são atribuídos a diferentes filas de acordo com suas características e o escalonador gerencia a CPU de acordo com as prioridades de cada fila.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

Além disso, o sistema operacional também é responsável por gerenciar a memória compartilhada entre os processos e garantir que cada processo tenha acesso apenas às áreas de memória alocadas a ele. Isso é feito por meio do uso de tabelas de páginas e do mapeamento de endereços virtuais para endereços físicos de memória.

O sistema operacional também pode usar técnicas como bloqueio de exclusão mútua e semáforos para garantir a sincronização e coordenação de acesso a recursos compartilhados pelos processos em execução no sistema multiprocessado.

X

## Redes de computadores

*conceitos básicos de redes de computadores, topologias de rede,  
protocolos de comunicação e arquiteturas de rede.*



## Conceitos básicos de redes de computadores

**R**edes de computadores são sistemas que permitem que computadores se comuniquem e compartilhem recursos, como dados, arquivos e dispositivos, como impressoras e scanners. A comunicação entre computadores em uma rede pode ocorrer por meio de conexões físicas, como cabos ou ondas de rádio, ou por meio de conexões lógicas, como a Internet.

Alguns dos conceitos básicos de redes de computadores incluem:

1. Topologia de rede: é a maneira como os dispositivos em uma rede estão conectados fisicamente. As topologias comuns incluem barramento, estrela e anel.
2. Protocolos de rede: são as regras e formatos que governam a comunicação entre dispositivos em uma rede. Os protocolos comuns incluem TCP/IP, HTTP, FTP e SMTP.
3. Endereço IP: é um número exclusivo que identifica um dispositivo em uma rede IP. É usado para rotear dados entre dispositivos em uma rede.
4. DNS: é um sistema de nome de domínio que converte nomes de domínio em endereços IP. Ele permite que os usuários acessem

sites usando nomes em vez de números IP.

5. Firewall: é um dispositivo ou software que monitora e controla o tráfego de rede para proteger uma rede contra ameaças externas.
6. VPN: é uma rede privada que usa a Internet pública para conectar dispositivos em locais geograficamente diferentes. Ele permite que os usuários accessem uma rede remotamente como se estivessem fisicamente presentes nela.
7. Servidores: são computadores que fornecem serviços, como armazenamento de arquivos, email e websites, para outros dispositivos em uma rede.
8. Clientes: são computadores ou dispositivos que acessam serviços fornecidos por servidores em uma rede.
9. Roteador: é um dispositivo que encaminha dados entre redes diferentes, permitindo que dispositivos em diferentes redes se comuniquem entre si.

Esses são apenas alguns dos conceitos básicos de redes de computadores. À medida que as redes se tornam mais complexas e avançadas, novos conceitos e tecnologias emergem para atender às necessidades dos usuários.

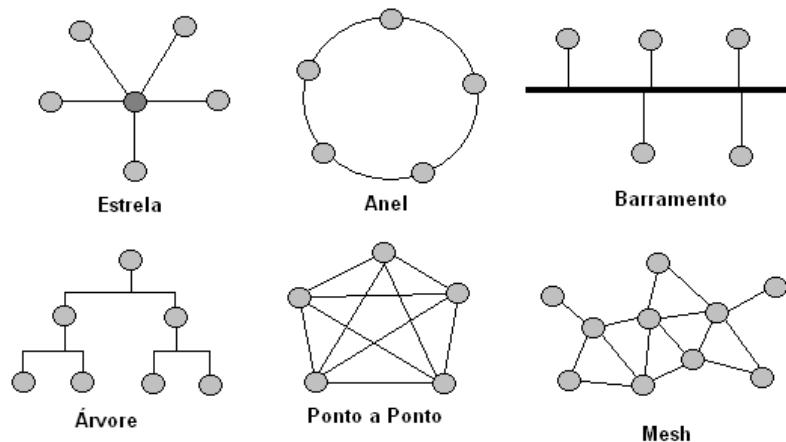
## Topologias de rede

**A**s topologias de rede se referem à forma como os dispositivos em uma rede são interconectados física e logicamente. Existem várias topologias de rede comuns, incluindo:

1. Topologia em estrela: todos os dispositivos são conectados a um único dispositivo central, geralmente um hub ou switch. É fácil de instalar e gerenciar, mas se o dispositivo central falhar, toda a rede pode ser afetada.
2. Topologia em barramento: todos os dispositivos são conectados a um único cabo de rede, que é compartilhado por todos os dispositivos. É fácil de instalar, mas se o cabo falhar, toda a rede pode ser afetada.
3. Topologia em anel: os dispositivos são conectados em uma configuração de anel, onde cada dispositivo está conectado a dois dispositivos vizinhos. É eficiente em termos de uso de cabo, mas se um dispositivo falhar, a rede pode ser interrompida.
4. Topologia em malha: todos os dispositivos são conectados uns aos outros em uma rede de malha. É muito redundante, o que significa que, se um dispositivo falhar, a rede ainda pode funcionar. No

entanto, pode ser caro e difícil de gerenciar.

Imagen das topologias:



Além dessas topologias, também existem topologias híbridas que combinam duas ou mais topologias. A escolha da topologia de rede dependerá das necessidades específicas de uma organização, incluindo a quantidade de dispositivos, a distância entre eles e o orçamento disponível.

## Protocolos de comunicação

Protocolos de comunicação são conjuntos de regras e procedimentos que definem como os dados são transmitidos entre dispositivos em uma rede de computadores. Eles garantem que a comunicação ocorra de maneira organizada, segura e eficiente.

Existem vários protocolos de comunicação, que são usados em diferentes camadas do modelo OSI (Open Systems Interconnection). Alguns exemplos de protocolos são:

- TCP/IP: é o protocolo mais usado em redes de computadores e é responsável por garantir a entrega confiável de dados entre dispositivos em uma rede. Ele é usado em várias camadas do modelo OSI, desde a camada de aplicação até a camada de transporte.
- HTTP: é o protocolo usado para transferência de dados na web. Ele é usado na camada de aplicação do modelo OSI e é responsável por solicitar e enviar páginas da web entre servidores e clientes.
- FTP: é o protocolo usado para transferência de arquivos entre dispositivos em uma rede. Ele também é usado na camada de aplicação do modelo OSI e é responsável por enviar e receber arquivos entre servidores e clientes.

## ARQUITETURA E ORGANIZAÇÃO DE COMPUTADORES, RESUMIDO EM UM LIVRO

- SMTP: é o protocolo usado para envio de e-mails. Ele é usado na camada de aplicação do modelo OSI e é responsável por enviar e receber mensagens de e-mail entre servidores e clientes.
- DNS: é o protocolo usado para resolução de nomes de domínio em endereços IP. Ele é usado na camada de aplicação do modelo OSI e é responsável por traduzir nomes de domínio em endereços IP para que os dispositivos possam se comunicar na rede.

Cada protocolo tem sua própria função e especificação técnica, mas todos são importantes para garantir a comunicação eficiente entre dispositivos em uma rede de computadores.

## Arquiteturas de rede

**A**s arquiteturas de rede referem-se ao conjunto de tecnologias, padrões e protocolos que definem a forma como os dispositivos de rede estão conectados e como eles se comunicam entre si. Existem várias arquiteturas de rede, cada uma com suas próprias vantagens e desvantagens. Algumas das arquiteturas de rede mais comuns incluem:

1. Rede em barramento (Bus): Nessa arquitetura, todos os dispositivos de rede são conectados a um único barramento central. Essa topologia é simples e barata, mas pode ser problemática quando há muitos dispositivos conectados e pode ocorrer congestionamento na rede.
2. Rede em anel (Ring): Nessa arquitetura, os dispositivos de rede são conectados em uma topologia de anel, em que cada dispositivo é conectado diretamente aos seus vizinhos. A comunicação ocorre no sentido do anel e cada dispositivo precisa esperar sua vez para enviar dados. Essa arquitetura é mais confiável do que a rede em barramento, mas pode ser lenta quando muitos dispositivos estão conectados.

3. Rede em estrela (Star): Nessa arquitetura, cada dispositivo de rede é conectado a um hub ou switch central, que gerencia o tráfego de dados na rede. Essa arquitetura é mais escalável do que as redes em barramento e em anel, mas é mais cara devido ao custo dos hubs ou switches.
4. Rede em malha (Mesh): Nessa arquitetura, cada dispositivo é conectado a vários outros dispositivos em uma rede de malha. Essa arquitetura é muito confiável e escalável, pois há várias rotas possíveis para os dados se movimentarem, mas é cara de implementar e gerenciar.
5. Rede sem fio (Wireless): Nessa arquitetura, os dispositivos de rede se conectam sem fios por meio de um ponto de acesso central. Essa arquitetura é muito flexível e fácil de implementar, mas pode ser menos confiável do que as redes com fio devido a interferências e limitações de alcance.

Essas são apenas algumas das arquiteturas de rede disponíveis, e cada uma pode ser adequada para diferentes cenários e requisitos de rede. É importante escolher a arquitetura de rede correta com base nas necessidades da organização.

## Also by Kauã de Jesus

Para escrever um livro sobre arquitetura e organização de computadores, precisei entender e explicar conceitos complexos de forma clara e objetiva, tornando-os acessíveis para estudantes e profissionais da área. Isso envolveu realizar uma pesquisa aprofundada, selecionar informações relevantes, organizar o conteúdo em capítulos e seções, e criar exemplos e exercícios práticos para auxiliar na compreensão dos leitores. Além disso, é necessário estar atualizado sobre as novas tendências e tecnologias na área, para que o livro seja útil e relevante por um longo período de tempo. O grande desafio foi conciliar a profundidade dos conceitos com a simplicidade da apresentação, para tornar o livro acessível e interessante para um público amplo.

Obrigado por ter chegado até aqui! Este livro contém um conteúdo resumido, mas espero que tenham aprendido muito e que tenha sido uma ótima leitura.



### **Arquitetura e organização de computadores para iniciantes**

Um livro com conhecimento resumido sobre a arquitetura e organização dos computadores

