

# Capítulo 1

## INTRODUÇÃO ÀS REDES DE COMPUTADORES

---



### 1. HISTÓRICO DAS REDES DE COMUNICAÇÃO

A evolução da microeletrônica e da informática tem possibilitado a obtenção de processadores e outros componentes de computadores cada vez mais potentes e velozes, num tamanho mais reduzido e num preço cada vez mais acessível a um maior número de pessoas.

Os microprocessadores existentes hoje em dia e que ocupam o espaço menor do que uma caixa de fósforos substituem e ultrapassam as capacidades dos computadores de alguns anos atrás, que ocupavam salas inteiras. Estes eram máquinas bastante complexas no que diz respeito à sua utilização e que ficavam em salas isoladas onde muito poucas pessoas tinham acesso, sendo operadas apenas por especialistas (analistas de sistema). Os usuários daqueles computadores normalmente submetiam seus programas aplicativos como *jobs* (ou tarefas) que executavam sem qualquer interação com o processamento do programa.

Uma primeira tentativa de interação com o computador ocorreu no início dos anos 60, com a técnica de *time-sharing*, que foi o resultado do desenvolvimento das teleimpressoras e da tecnologia de transmissão de dados. Nesta técnica um conjunto de terminais era conectado a um computador central através de linhas de comunicação de baixa velocidade, o que permitia aos usuários interagir com os seus programas. A necessidade de conexão de terminais para o processamento interativo foi o ponto de partida para o estabelecimento de necessidades de comunicação nos computadores. A técnica de *time-sharing* permitia a um grande conjunto de usuários o compartilhamento de um único computador para a resolução de uma grande diversidade de problemas e as aplicações desenvolvidas foram cada vez mais se multiplicando e se diversificando (cálculos complexos, produção de relatórios, ensino de programação, aplicações militares, etc...). Este aumento na demanda implicava numa necessidade crescente de atualizações e incrementos nas capacidades de cálculo e de armazenamento nas unidades central o que nem sempre era viável ou possível, dado que os computadores do tipo *mainframes* nem sempre eram adaptados para suportar determinadas extensões.

Nos anos 70, com o surgimento dos minicomputadores, foi possível adaptar as capacidades de processamento às reais necessidades de uma dada aplicação. Além disso, dado que a nível de uma empresa um grande número de usuários operavam sobre conjuntos comuns de informações, a necessidade do compartilhamento de dados, de dispositivos de armazenamento e de periféricos entre os vários departamentos de uma empresa deu um novo impulso aos trabalhos no sentido de se resolver os problemas de comunicação entre os computadores. Este novo tipo de aplicações exigia uma velocidade e uma capacidade de transmissão muito mais elevadas que no caso da conexão de terminais a um computador central. Assim, com a utilização de minicomputadores interconectados, obtinha-se uma capacidade de processamento superior àquela possível com a utilização dos *mainframes*. Outro aspecto interessante é que as redes podiam ser estendidas em

função das necessidades de processamento das aplicações. Além disso, a modularidade natural das redes de computadores era tal que uma falha num minicomputador (ou de comunicação da rede) tinha um efeito bastante limitado em relação ao processamento global.

Atualmente, as vantagens dos sistemas distribuídos e interconectados são uma evidência reconhecida para as aplicações mais diversas, desde a automação de escritórios até o controle de processos, passando por aplicações de gerenciamento bancário, reservas de passagens aéreas, processamento de texto, correio eletrônico, etc...

## 2. IMPORTÂNCIA DAS REDES DE COMUNICAÇÃO

Um grande número de empresas possui atualmente uma quantidade relativamente grande de computadores operando nos seus diversos setores. Um exemplo deste fato é aquele de uma empresa que possui diversas fábricas contendo cada uma um computador responsável das atividades de base da fábrica (controle de estoques, controle da produção e, o que também é importante, a produção da folha de pagamentos). Neste exemplo, apesar da possibilidade de operação destes computadores de maneira isolada, é evidente que sua operação seria mais eficiente se eles fossem conectados para, por exemplo, permitir o tratamento das informações de todas as fábricas da empresa. O objetivo da conexão dos diferentes computadores da empresa é permitir o que poderíamos chamar de *compartilhamento de recursos*, ou seja, tornar acessíveis a cada computador todos os dados gerados nas diversas fábricas da empresa.

Um outro ponto importante da existência das Redes de Comunicação é relacionado a um aumento na *confiabilidade* do sistema como um todo. Pode-se, por exemplo, ter multiplicados os arquivos em duas ou mais máquinas para que, em caso de defeito de uma máquina, cópias dos arquivos continuarão acessíveis em outras máquinas. Além disso, o sistema pode operar em regime *degradado* no caso de pane de um computador, sendo que outra máquina pode assumir a sua tarefa. A continuidade de funcionamento de um sistema é ponto importante para um grande número de aplicações, como por exemplo: aplicações militares, bancárias, o controle de tráfego aéreo, etc.

A *redução de custos* é uma outra questão importante da utilização das Redes de Comunicação, uma vez que computadores de pequeno porte apresentam uma menor relação preço/desempenho que os grandes. Assim, sistemas que utilizariam apenas uma máquina de grande porte e de custo muito elevado podem ser concebidos à base da utilização de um grande número de microcomputadores (ou estações de trabalho) manipulando dados presentes num ou mais servidores de arquivos.

## 3. EXTENSÃO E TOPOLOGIA

### 3.1. Redes locais e redes de longa distância

Na seção anterior foram apresentados dois exemplos de implementação de Redes de Comunicação: no primeiro caso, o sistema era composto de diversos computadores espalhados cada um numa fábrica da empresa.

No segundo caso, o sistema era composto de diversos microcomputadores, podendo todos estar localizados na mesma sala ou em salas vizinhas num mesmo edifício.

A diferença na dimensão das Redes de Comunicação introduz diferentes problemas e necessidades e deve então, fazer objeto de uma classificação. No que diz respeito ao exemplo dos microcomputadores, a rede é classificada como sendo uma *Rede Local* (ou *LAN - Local Área Network*), caracterizada particularmente por uma pequena extensão, limitando-se normalmente à interconexão de computadores localizados numa mesma sala, num mesmo prédio ou num campus.

No exemplo da empresa possuindo diversas fábricas, a rede utilizada permitiria conectar computadores localizados em diferentes prédios numa mesma cidade ou mesmo

em cidades distantes de uma dada região. Esta caracteriza uma *Rede de Longa Distância* ou *Rede Geograficamente Distribuída* (ou WAN - Wide Área Network).

### 3.2. As diferentes topologias

Um ponto importante no que diz respeito à concepção de uma rede de comunicação é a definição da maneira como as diferentes estações serão associadas. Inicialmente, podemos distinguir dois tipos principais de concepção: os canais em modo ponto-a-ponto e os canais de difusão.

Nos *canais em ponto-a-ponto*, a rede é composta de diversas linhas de comunicação, cada linha sendo associada à conexão de um par de estações.

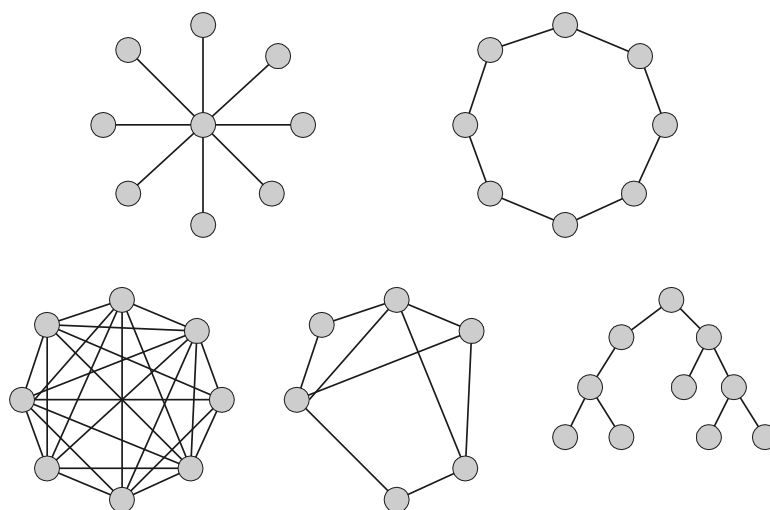
Neste caso, se duas estações devem se comunicar sem o compartilhamento de um cabo, a comunicação será feita de modo indireto, através de uma terceira estação. Assim, quando uma mensagem (ou *pacote*) é enviada de uma estação a outra de forma indireta (ou seja, através de uma ou mais estações), ela será recebida integralmente por cada estação e, uma vez que a linha de saída da estação considerada está livre, retransmitida à estação seguinte.

Esta política de transmissão é também conhecida por “*store and forward*” ou comutação de pacotes. A maior parte das redes de longa distância são do tipo ponto-a-ponto.

As redes ponto-a-ponto podem ser concebidas segundo diferentes topologias. As redes locais ponto-a-ponto são caracterizadas normalmente por uma topologia simétrica; as redes de longa distância apresentam geralmente topologias assimétricas. A figura 1.1 apresenta as diferentes topologias possíveis nas redes ponto-a-ponto.

Uma outra classe de redes, as *redes de difusão*, são caracterizadas pelo compartilhamento, por todas as estações, de uma linha única de comunicação. Neste caso, as mensagens enviadas por uma estação são recebidas por todas as demais conectadas ao suporte, sendo que um campo de endereço contido na mensagem permite identificar o destinatário.

Na recepção, a máquina verifica se o endereço definido no campo corresponde ao seu e, em caso negativo, a mensagem é ignorada. As redes locais pertencem geralmente a esta classe de redes. Nas redes de difusão, existe a possibilidade de uma estação enviar uma mesma mensagem às demais estações da rede, utilizando um código de endereço especial; neste caso, todas as estações vão tratar a mensagem recebida.



**Figura 1.1** - Topologias ponto-a-ponto: estrela, anel, malha regular, malha irregular e árvore.

Pode-se ainda especificar uma mensagem de modo que esta seja enviada a um subgrupo de estações da rede. A figura 1.2 apresenta algumas topologias possíveis no caso das redes a difusão. Numa rede em barramento, uma única máquina é autorizada a cada

instante a transmitir uma mensagem — é a estação *mestre* do barramento. As demais estações devem esperar autorização para transmissão.

Para isto, um mecanismo de *arbitragem* deve ser implementado para resolver possíveis problemas de conflito (quando duas ou mais estações querem enviar uma mensagem), este mecanismo podendo ser centralizado ou distribuído.

No caso das redes de satélite (ou rádio), cada estação é dotada de uma antena através da qual pode enviar e receber mensagens. Cada estação pode “escutar” o satélite e, em alguns casos, receber diretamente as mensagens enviadas pelas demais estações.

No caso do anel, cada bit transmitido é propagado de maneira independente em relação à mensagem (ou pacote) ao qual ele pertence. Em geral, cada bit realiza uma volta completa do anel durante o tempo necessário para a emissão de um certo número de bits, antes mesmo da emissão completa da mensagem.

Também nesta topologia, é necessária a implementação de um mecanismo de acesso ao suporte de comunicação. Existem diferentes técnicas para este fim que serão discutidas ao longo do curso.

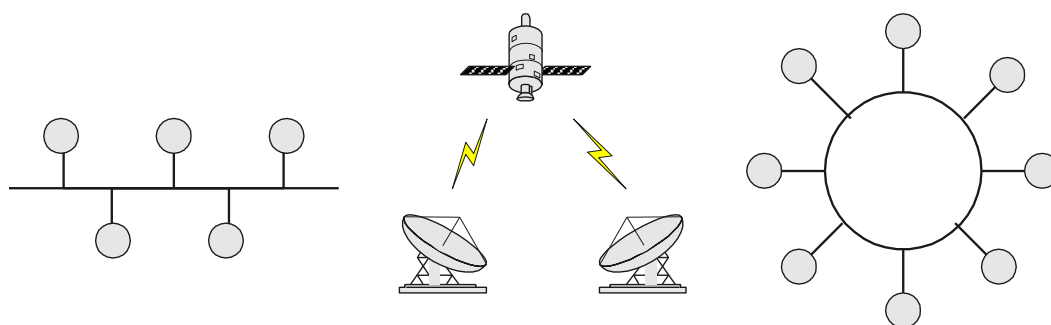
As redes de difusão podem ainda considerar duas classes de mecanismos de acesso ao suporte de comunicação: estáticas ou dinâmicas. Um exemplo do primeiro caso é a definição de intervalos de tempo durante os quais cada estação tem a posse do canal de comunicação, permitindo então que esta emita a mensagem de maneira cíclica. No entanto, esta política é bastante ineficiente do ponto de vista do envio das mensagens, uma vez que muitas estações não vão enviar mensagens nos intervalos a elas destinadas. Já na outra classe de mecanismos, dinâmicos, o acesso é dado às estações segundo a demanda de envio de mensagens. Nos mecanismos de acesso dinâmicos, pode-se ainda considerar dois casos:

- os mecanismos centralizados, nos quais uma estação central (árbitro) é a responsável da definição do direito de acesso ao suporte de comunicação;
- os mecanismos distribuídos, nos quais cada estação define quando ela vai emitir a mensagem.

## 4. ASPECTOS ARQUITETURAIS

### 4.1. Serviços necessários à comunicação

Como visto nas seções precedentes, as redes de computadores podem se caracterizar por diferentes configurações e topologias. Apesar da diversidade no que diz respeito a este aspecto, todas as possíveis configurações têm um objetivo comum — a *transferência de dados*.



**Figura 1.2** - Topologias das redes de difusão: barramento, satélite e anel

O problema que se coloca é então relacionado à especificação dos procedimentos e mecanismos que devem ser implementados para viabilizar o funcionamento da rede. A resolução deste problema é baseada principalmente no conhecimento prévio das funções que devem ser suportadas pela rede, assim como do ambiente no qual ela vai ser inserida.

Estes aspectos serão mostrados aqui através de alguns exemplos. O primeiro exemplo é baseado nas políticas de *time-sharing* já descrita na seção 1. Vamos considerar o caso em que temos apenas um terminal conectado a um computador, como mostrado na figura 1.3.

Considerando que um usuário vai servir-se do terminal para processar informações no computador central, para que isto seja possível, é necessário que o computador central seja dotado do programa necessário ao tratamento daquelas informações. Em caso positivo, o terminal e o computador devem estabelecer um diálogo que permita o bom desenrolar das operações de tratamento das informações. Este diálogo deverá permitir, por exemplo, que o usuário comunique sua intenção (de processar as informações!) ao computador e, em seguida, envie as informações a serem processadas. Uma vez efetuado o tratamento, o computador deve retornar os resultados ao terminal.

Esta seqüência de operações, apesar de aparentemente elementar, requer a satisfação de uma série de condições. Vamos supor, por exemplo, que o computador central e o terminal tenham sido fabricados de forma totalmente independente um do outro, o que pode ter resultado numa diferente filosofia no que diz respeito ao formato das informações.

Um primeiro obstáculo a ser vencido é aquele da linguagem; o terminal deveria então se adaptar à linguagem do computador central.

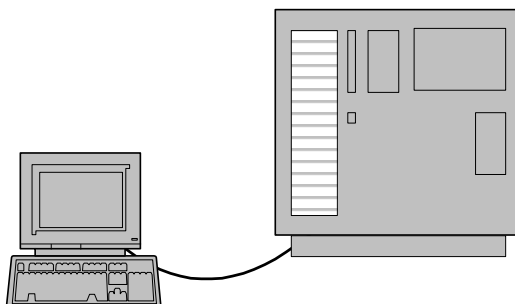
Resolvido o problema de compreensão, um outro problema encontrado diz respeito aos possíveis erros de transmissão que podem ocorrer durante a comunicação, uma vez que as linhas de comunicação estão sujeitas a ruídos e outros fenômenos podendo provocar perdas de informação.

Uma outra questão pode ainda estar relacionada à velocidade de funcionamento dos dois elementos. Se considerarmos que o computador central opera numa velocidade superior à do terminal, por exemplo, o terminal corre o risco de ser “bombardeado” pelo fluxo de dados vindo do computador, o que vai exigir então o estabelecimento de um mecanismo de controle do fluxo de informação.

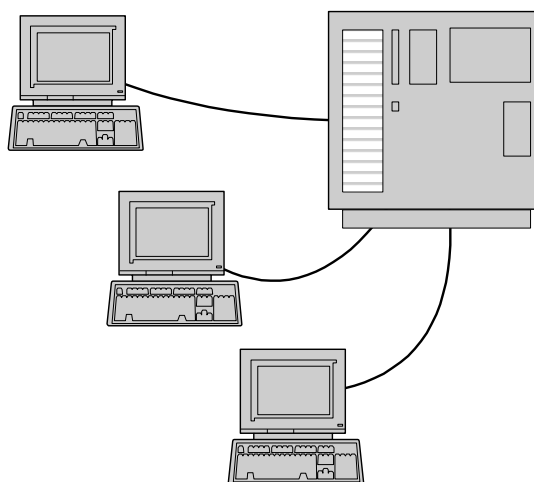
Resumindo, a rede de comunicação deve, além de suprir as funções de transmissão e tratamento de informações, oferecer serviços de adaptação, detecção e correção de erros de transmissão e controle de fluxo.

Vamos considerar agora que, ao invés de um único terminal, vamos conectar um maior número deles ao computador central (figura 1.4).

Aqui, cada terminal pode, a princípio e a qualquer momento, tomar a iniciativa da troca de dados com o computador. Isto significa que cada terminal terá de ser caracterizado por um endereço específico, cuja utilização correta vai permitir evitar que o computador central envie as informações aos terminais de maneira indevida.



**Figura 1.3** - Terminal conectado a um computador central



**Figura 1.4 - Configuração com vários terminais**

Por outro lado, se o número de terminais conectados ao computador central torna-se relativamente elevado (a fim de permitir a utilização máxima da capacidade de processamento deste), será necessário organizar as interações entre terminais e o computador central em sessões, de tal forma que, ao término de uma sessão entre um terminal e o computador central, este terá liberados determinados elementos (envolvidos naquela sessão) que poderão atender outros terminais em estado de espera.

Ainda, considerando que nem todos os terminais vão efetuar o mesmo tipo de tratamento de forma simultânea, dever-se-á, então, especificar a aplicação associada. Assim, todas as necessidades vistas neste exemplo deverão ser associadas às funcionalidades definidas no exemplo anterior. Mas os problemas não terminam por aqui... (é impossível, no momento, prever onde terminarão os problemas!)

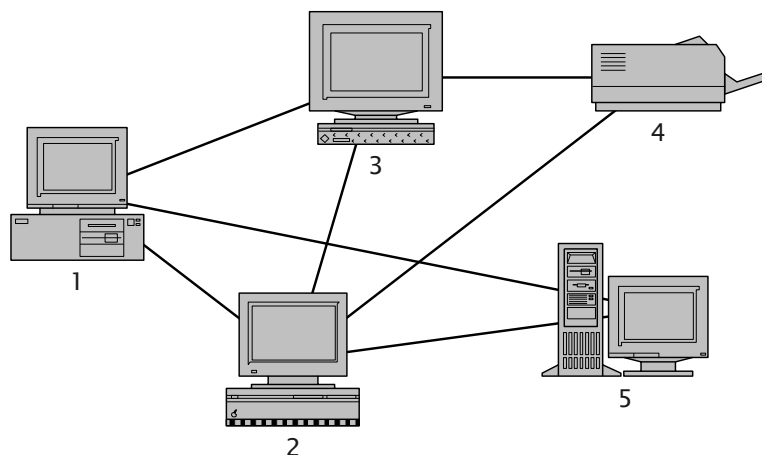
Vamos considerar ainda um exemplo, mais particularmente o de uma rede contendo diversos computadores, terminais, etc... , cada um destes elementos constituindo um nó da rede (figura 1.5). Neste exemplo, os dois elementos envolvidos numa comunicação não serão mais necessariamente adjacentes; além disso, podem existir diversas maneiras de conectá-los, o que vai corresponder a diferentes caminhos. No exemplo mostrado na figura, os nós 1 e 5 podem ser conectados por pelos menos uns 10 caminhos e a escolha de qual caminho utilizar deverá então ser realizada, o que é não é uma tarefa tão simples quanto possa parecer.

Ainda, se a rede é da classe ponto-a-ponto (ou comutação de pacotes), o sistema de comunicação deve assegurar a correta transmissão (transporte) da informação de um ponto a outro. Particularmente, será necessário garantir que as mensagens enviadas serão recuperadas e reconstituídas na ordem correta no ponto de chegada.

Um requisito também importante é o aspecto da codificação das mensagens de modo a evitar o acesso a informações de parte de usuários alheios ao sistema considerado. A esta função, pode-se eventualmente acrescentar a técnicas de compressão de dados, necessária se a informação enviada é demasiadamente redundante e o custo da comunicação é alto.

## 4.2. Questões organizacionais

Uma vez listadas as diferentes necessidades relacionadas a uma rede de comunicação, a questão que se coloca é a da viabilidade de um projeto de rede, dada a quantidade de funções a implementar.

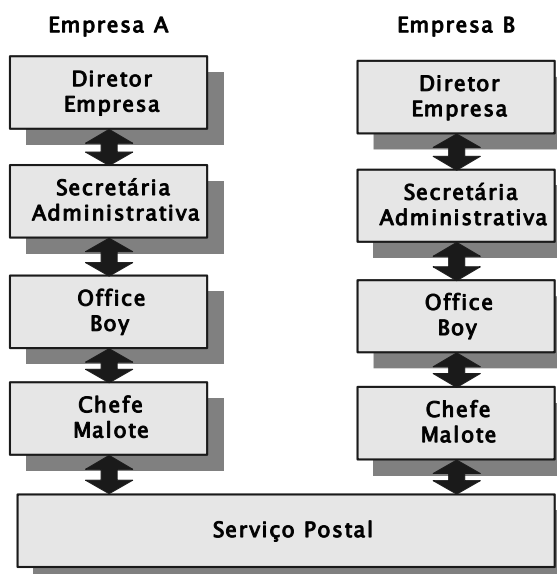


**Figura 1.5** - Configuração com vários computadores e terminais.

Uma outra questão é a do ordenamento das funções. O controle de fluxo deve ser realizado antes ou depois da correção de erros? Uma vez resolvida esta questão, que elementos da rede serão responsáveis da implementação destas funções? As soluções adotadas são dependentes do suporte de transmissão utilizado? Elas continuam válidas no caso de expansão da rede? Estas questões representam, de certo modo, a necessidade de levar em conta um certo ordenamento no que diz respeito à adoção das soluções a cada problema.

Uma ilustração típica do problema é aquela da comunicação entre duas empresas. Vamos supor que o Diretor de uma Empresa A quer comunicar com o Diretor de uma Empresa B. Ele convoca a sua Secretária Administrativa e solicita, informalmente, que esta construa um texto relativo ao assunto a ser tratado. A Secretária Administrativa elabora o documento e o entrega ao *Office Boy* que vai envelopá-lo e encaminhá-lo ao Chefe do Setor de Malote. Este último encaminha o documento ao Serviço Postal para condução à Empresa B.

Considerando que a Empresa B apresenta uma estrutura similar à Empresa A (como é ilustrado na figura 1.6), os mesmos elementos atuam, cada um em suas funções para fazer com que a correspondência chegue às mãos do Diretor da Empresa.



**Figura 1.6** - Filosofia de concepção das redes, ilustrada por um processo de relações entre empresas.

Este processo caracteriza, na verdade, a filosofia de concepção das redes de comunicação, que é baseada em dois conceitos fundamentais: o da hierarquia e o da descentralização, cuja conjunção vai permitir responder à questão de ordenação na adoção das soluções. Segundo esta filosofia, uma tarefa global é vista como sendo decomposta à medida que se vai descendo na hierarquia e que a única interação física se faz no seu nível mais baixo.

Podemos considerar que a comunicação entre dois nós de uma rede é uma tarefa global que afeta um sistema complexo e, conseqüentemente, sujeita à aplicação dos princípios de hierarquização e de descentralização.

As vantagens da adoção destes princípios são, fundamentalmente:

- facilidade de estudo e de implementação da rede a partir de elementos de base existentes, o que permite a redução dos custos de instalação;
- simplificação de sua operação em função da definição de regras formais;
- garantia de confiabilidade de um sistema que seja aceitável, particularmente graças ao encapsulamento das funções o que permite limitar a propagação de erros e facilitar a manutenção;
- garantia, pela modularidade, de um grau satisfatório de evolutividade e de extensibilidade da rede;
- otimizar o desempenho.

Todos estes aspectos nos conduzem a conceber uma arquitetura de comunicação como sendo uma organização de software e hardware estruturada em camadas.

#### 4.3. A estruturação em camadas

Os conceitos de hierarquia e descentralização podem ser empregados de diferentes formas, cada uma podendo implicar num tipo de rede particular. Em função desta provável multiplicidade, surgiu então a necessidade de uma normalização permitindo a conexão de diferentes classes de hardware.

Para possibilitar a normalização, foi necessário estabelecer um modelo teórico capaz de representar as relações entre as diferentes tarefas implementadas nos diferentes níveis hierárquicos. A possibilidade de interconexão de um número qualquer de sistemas, ou seja, de conjuntos autônomos podendo efetuar tarefas de tratamento ou de transmissão de informação, era uma característica essencial para o modelo a ser estabelecido.

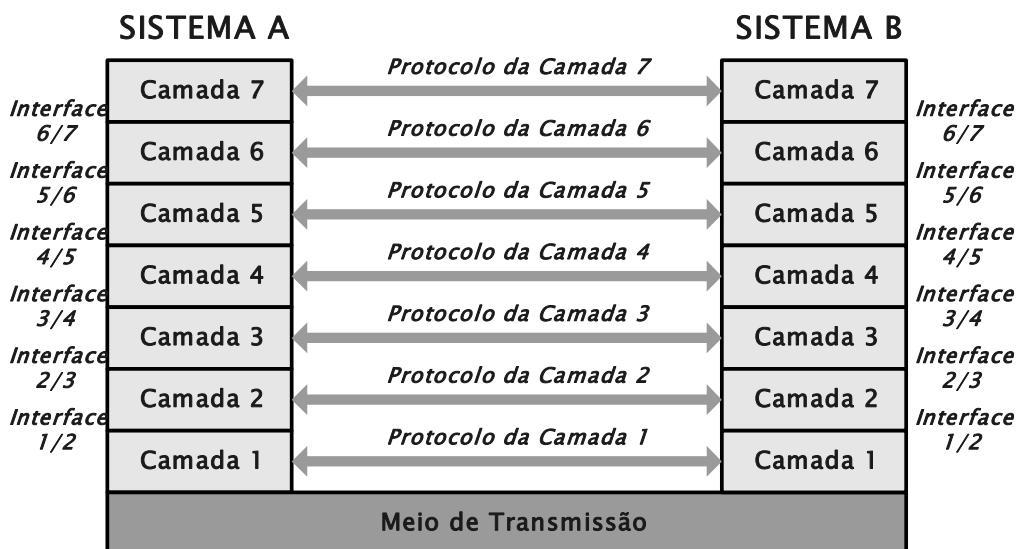
A figura 1.7 ilustra uma arquitetura hierarquizada em 7 camadas que permitirá introduzir o conjunto de conceitos relacionados ao modelo estabelecido.

O objetivo de cada camada é o oferecimento de um determinado serviço às camadas superiores de forma a evitar que estas necessitem conhecer certos aspectos da implementação destes serviços.

A camada  $n$  assume a comunicação com a camada  $n$  de uma outra máquina. Para fazê-lo, ela se serve de um conjunto de convenções e regras que vão permitir gerir esta comunicação. A este conjunto de regras e convenções, dá-se o nome de *protocolo* da camada  $n$ , ou, simplesmente, *protocolo  $n$* . Como se pode ver na figura, não existe meio de comunicação físico entre as diferentes camadas (apenas o Meio de Transmissão entre as entidades pares da camada 1), o que significa que não existe transferência direta de dados entre a camada  $n$  de uma máquina à camada  $n$  de outra máquina.

Na realidade, cada camada transfere os dados à camada imediatamente inferior até a camada mais baixa; o dado é então transmitido à outra máquina através do Meio de Transmissão. A comunicação entre as camadas é vista então como uma comunicação virtual e é representada, na figura 1.7, pelas linhas tracejadas ligando cada par de processos de uma camada.





**Figura 1.7** - Modelo hierarquizado em 7 camadas.

Cada camada comunica-se com as camadas adjacentes através de uma interface, que define as operações elementares e os serviços que a camada inferior oferece à camada considerada.

No momento da definição do número de camadas que vai compor uma rede e do papel que cada uma delas deve cumprir, uma tarefa importante será a definição completa das interfaces entre as camadas; isto vai implicar que na definição do serviço oferecido por cada camada.

Uma vantagem da correta definição das interfaces é a facilidade da introdução de modificações nas implementações das diferentes camadas; os mecanismos podem ser implementados de forma diferente, desde que as interfaces anteriormente definidas sejam respeitadas.

Ao conjunto das camadas compondo uma rede dá-se o nome de *arquitetura da rede*, e as especificações da arquitetura devem conter informações suficientes para permitir o correto desenvolvimento da rede, tanto do ponto de vista do software quanto do hardware. Por outro lado, os detalhes de implementação dos mecanismos a implementar em cada camada, assim como as especificações detalhadas das interfaces não fazem parte da definição da arquitetura da rede.

## 5. TECNOLOGIAS DE TRANSMISSÃO DE DADOS

### 5.1. Principais Características

Podemos definir transmissão como a técnica do transporte do sinal por um meio, de um ponto a outro afastado. Em particular, a transmissão de dados apresenta diversas características referentes ao sentido da transmissão, número de canais utilizados, sincronismo entre transmissor e receptor e velocidade de transmissão.

Um equipamento pode ser projetado de tal forma que a transmissão sobre um determinado meio seja feita em uma das seguintes formas (sentido):

- **Simplex** - quando a transmissão é feita em um único sentido.
- **Half-duplex** - quando a transmissão é feita nos dois sentidos mas não ao mesmo tempo.
- **Full-duplex** - quando a transmissão é feita nos dois sentidos simultaneamente.

Uma mensagem é definida como um conjunto de símbolos. Cada símbolo, por sua vez, para efeito de transmissão de dados, é caracterizado por um conjunto de configurações do sinal que representam *bits*. Por necessidade de codificação, os símbolos ficam associados a caracteres, que são, na realidade, configurações de sinais. Na verdade, uma mensagem nada mais é que uma seqüência de *bits*. Para transferir essa seqüência de *bits*, podemos fazer de duas formas: **serial** ou **paralela**. Na transmissão paralela, os *bits* que compõem um carácter são transportados de forma simultânea, cada um possuindo seu próprio canal, conforme mostra a figura 1.8.

Na transmissão serial, os *bits* que compõem um carácter são transportados um após o outro, utilizando apenas um canal, como ilustrado na figura 1.9.

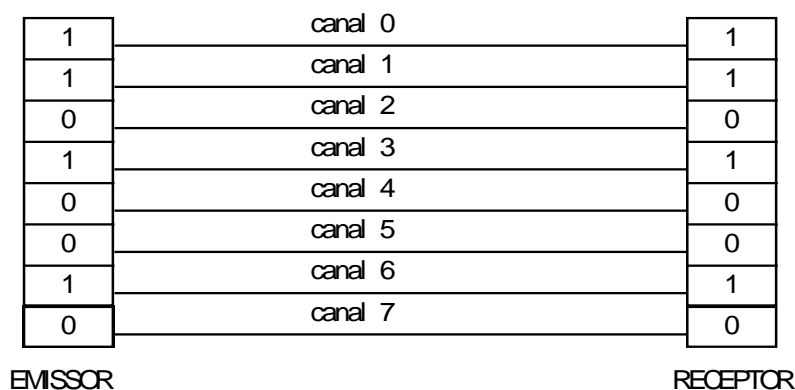


Figura 1.8 - Transmissão paralela.

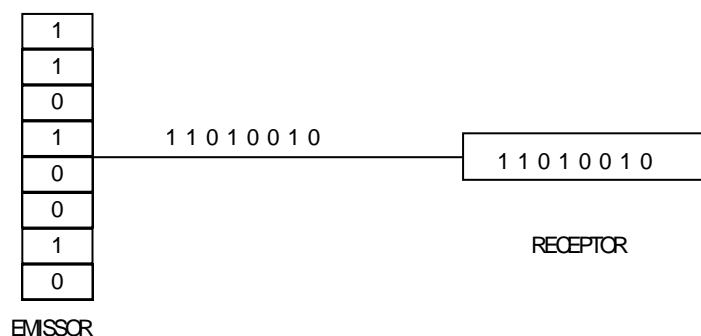


Figura 1.9 - Transmissão serial.

Como os *bits* chegam um de cada vez, o equipamento receptor deverá saber qual *bit* é o primeiro do carácter, a fim de que possa decodificar o símbolo recebido, ou quais *bits* são realmente de informação. Este é um problema de **sincronização**. Sincronização pode ser vista, então, como o método do equipamento transmissor fazer a separação dos caracteres ou das mensagens para o equipamento receptor. Existem dois modos de se fazer sincronização e eles são conhecidos como **transmissão síncrona** e **transmissão assíncrona**. Todos os dois modos dizem respeito à transmissão serial.

Na transmissão serial assíncrona, a sincronização é alcançada precedendo-se cada carácter de um elemento de *start* e colocando-se, após os *bits* de dados um elemento de *stop*. Geralmente os dados são transmitidos serialmente em um número fixo de *bits*, dependendo do código usado.

Na transmissão serial síncrona, a sincronização é alcançada transmitindo-se os *bits* de um carácter seguidos imediatamente pelos bits do próximo carácter, não havendo elementos de *start/stop* entre eles. O conjunto de caracteres que formam uma mensagem é dividido em blocos. O tamanho destes blocos varia desde alguns poucos caracteres até centenas deles, dependendo dos equipamentos envolvidos na transmissão.

A sincronização na transmissão síncrona se faz antes que um bloco de informação seja enviado pelo equipamento transmissor. Ela é alcançada pelo envio de uma configuração de *bits* chamada "caracter de sincronização" que precede o bloco a ser transmitido. Neste caso, a sincronização é feita mediante um ou dois bytes de sincronismo que são configurações de *bits* designadas para este fim. Do mesmo modo que é iniciada, a comunicação síncrona é encerrada, não por um *bit*, mas por um ou mais caracteres de fim de transmissão.

Os caracteres iniciais e finais de um bloco em transmissão síncrona constituem o envelope do bloco, que pode ou não ser uma mensagem completa e tem funções análogas as do envelope de uma carta, inclusive a de conter também o endereço do destinatário. A ordem correta dos *bits* que constituem o envelope do bloco é definida por um conjunto de regras chamado **protocolo de comunicação**. São exemplos destes protocolos para transmissão síncrona: BSC, SDLC e HDLC.

As vantagens de transmissão assíncrona são: geração de caracteres por meio de dispositivos eletromecânicos e transmissão de caracteres irregularmente espaçados no tempo. As desvantagens são: uma parte considerável do que transmite não transporta informação útil: a sincronização depende dos *start/stop* que podem não ser detectados por causa de distorções do sinal. Uma utilização da transmissão assíncrona é quando não se necessita de transmissão freqüente de informação.

A transmissão síncrona apresenta, como vantagem, uma melhor proteção contra erros, pois ao término de cada bloco uma configuração de *bits* para detecção de erros pode ser enviada; é mais eficiente pois a proporção de mensagem transmitida como informação em relação à configuração de sincronização é maior que na transmissão assíncrona; não é tão sensível à distorção e opera a velocidades bem mais altas que no modo assíncrono. As desvantagens são: caso haja erro de sincronização, todo bloco é perdido; os caracteres são enviados em blocos e não antes destes poderem ser formados, obrigando que os equipamentos sejam dotados de memória de armazenamento para a coleta dos caracteres até que se forme o bloco com o comprimento usado pelo equipamento. Memória, nesse caso, são *buffers*, o que encarece seu custo.

## 5.2. Largura de Banda e Capacidade de Canal

A taxa em que podemos enviar dados sobre um canal é proporcional à **largura de banda** do canal (ou *bandwidth*). O termo largura de banda não tem qualquer relação com as freqüências que são transmitidas no canal. Ele indica apenas a diferença entre os limites inferior e superior das freqüências que são suportadas pelo canal. Por exemplo, um canal que admite freqüências da ordem de 1500 a 5000 Hz, tem uma largura de banda igual a  $5000 - 1500 = 3500\text{Hz}$ . Da mesma forma, um canal que admite freqüências que vão desde 18000 Hz a 21500 Hz também apresenta uma largura de banda 3500 Hz ( $21500 - 18000$ ).

A largura de banda de um canal de comunicação constitui uma medida da máxima taxa de informação que pode ser transmitida pelo canal. Largura de banda significa o espectro de freqüência que o canal é capaz de transmitir. Sobre a questão de quantos estados de sinalização podem ser transmitidos e distinguidos separadamente no receptor de um sistema de comunicação de dados pode-se dizer que, examinados os fatores que influenciam esse número de estados, podemos definir o conceito de **capacidade do canal**.

Ruído, distorção e flutuações na atenuação do sinal portador têm influência no número de estados de sinalização. Este número de estados é observado na unidade de tempo (segundo). Quanto maior o número de estados que podem ser transmitidos e distinguidos, maior será a capacidade do canal. Podemos então concluir que a capacidade do canal está intimamente relacionada com a velocidade de transmissão, pois quanto maior o número de estados mais *bits* por segundo poderão ser transmitidos. Daí medir-se capacidade na unidade *bits/segundo*.

Diferentes tipos de sinais (voz humana, música, dados, imagem) necessitam de diferentes capacidades de canal, as quais são indicadas em termos de largura de banda e outros fatores que influenciam a capacidade de um canal. Conhecida, pois, a largura de

banda de um canal de comunicação, pode-se estabelecer a máxima taxa de sinalização que o mesmo pode conduzir sem erro, o que é denominado de capacidade do canal de comunicação.

### 5.3. Distorção e Ruído na Transmissão

**Atenuação, ruído e retardo** são termos usados de um modo geral para descrever as modificações que um sinal sofre quando é transmitido em um circuito ou canal. Essas alterações resultam de imperfeições na propagação do sinal, ao longo do suporte de transmissão (atenuação e retardo) e de perturbações (ruídos) que atuam não só no suporte de transmissão como também nos estágios de processamento do sinal que compõem o receptor. Quando a alteração sofrida pelo sinal é muito grande, pode ocorrer a detecção trocada da informação binária, isto é, detecção de "1" quando foi transmitido o "0" ou detecção de "0" quando foi transmitido o "1". Estas detecções trocadas caracterizam os chamados erros de transmissão.

O desempenho de um sistema de transmissão de dados é avaliado através do seu grau de confiabilidade na transmissão dos *bits*. A taxa de erros de um sistema de transmissão representa a probabilidade de ocorrência de erros de transmissão.

Para redes de longa distância existe uma padronização internacional que determina uma taxa de erros máxima em um canal a fim de que o mesmo possa ser considerado adequado para a transmissão de dados. O ITU recomenda uma taxa de erros não superior a  $10^{-5}$ , isto é, 1 *bit* errado a cada 100.000 *bits* transmitidos. Para redes locais, de maneira geral, admite-se taxas de erros típicas da ordem de  $10^{-9}$  a  $10^{-12}$ .

A transmissão sem erros é um requisito essencial de quase todas as aplicações de comunicação de dados e portanto, uma série de mecanismos deve ser implementada para detectar e corrigir possíveis erros. A maneira usual utilizada para detectar a alteração de *bits* de informação transmitidos, consiste na adição de *bits* de redundância na mensagem a ser transmitida. Dentre as várias técnicas usadas para esta finalidade pode-se citar: *bits* de paridade por carácter, paridade longitudinal e códigos de redundância cíclica (CRC).

### 5.4. Comutação

A função de **comutação**, ou **chaveamento**, em uma rede de comunicação refere-se à alocação dos recursos da rede (meios de transmissão, repetidores, sistemas intermediários, etc.) para a transmissão pelos diversos dispositivos conectados. A alocação destes recursos está intimamente ligada à forma de multiplexação dos meios de transmissão, que será discutida mais adiante. As formas de comutação são denominadas **comutação de circuitos**, **comutação de mensagens** e **comutação de pacotes**.

A comutação de circuitos pode ser realizada se existir um caminho dedicado de comunicação entre duas estações. Esta comunicação envolve 3 fases: estabelecimento do circuito, transferência de informações e desconexão do circuito. Na comutação de circuitos o caminho alocado durante a fase de estabelecimento da conexão permanece dedicado àquelas estações até que uma delas, ou ambas, decida desfazer o circuito. Caso o tráfego entre as estações não seja contínuo e constante a capacidade do meio físico será desperdiçada. O caminho dedicado entre a origem e o destino pode ser: um caminho físico formado por uma sucessão de enlaces físicos, uma sucessão de canais de frequência alocados em cada enlace ou uma sucessão de canais de tempo alocados em cada enlace. Este tipo de comutação é bastante utilizada em sistemas telefônicos.

Na comutação de mensagens não é necessário o estabelecimento de um caminho dedicado entre as estações. Se uma estação deseja transmitir uma mensagem ela adiciona o endereço de destino da mesma poderá ser transmitida pela rede de nó em nó, utilizando apenas um canal por vez. A mensagem é armazenada e retransmitida em cada nó (*store-and-forward*).

A comutação de pacotes é semelhante à comutação de mensagens, sendo que a principal diferença está no tamanho da mensagem sendo transmitida, que neste caso tem

tamanho reduzido e limitado. Neste caso as mensagens com tamanho acima de um limite devem ser fragmentadas em unidades menores denominadas **pacotes**. Os pacotes que formam uma mensagem podem estar sendo transmitidos simultaneamente pela rede em diferentes enlaces, o que reduz o atraso de transmissão total da mensagem. Redes com comutação de pacotes requerem nós de comutação com menor capacidade de armazenamento e procedimentos de recuperação de erros mais eficientes do que para comutação de mensagens.

## 5.5. Multiplexação

Quando a banda passante de um meio físico for maior ou igual à banda passante necessária para um sinal podemos transmitir mais de um sinal neste meio físico. Esta técnica é denominada **multiplexação**.

As duas principais técnicas de multiplexação são: **multiplexação por divisão de frequência** (FDM) e **multiplexação por divisão de tempo** (TDM). A primeira técnica divide a largura de banda em diversas larguras de banda menores, cada qual com seu próprio domínio de frequências. Cada uma destas componentes pode ser usada individualmente como se ela fosse uma linha separada. A segunda técnica intercala os *bits*, que fluem das linhas de baixa velocidade, dentro da linha de maior velocidade. Em ambos os métodos o resultado é que uma linha transmite em paralelo um número de sinais de linhas de velocidades mais baixas.

Uma desvantagem da FDM é a relativa dificuldade de expansão porque, uma vez determinado o número de sub-canais sobre a linha compartilhada, pode ser difícil adicionar novos sub-canais já que as frequências em um grupo devem ser reatribuídas. Na TDM na qual o tempo disponível é dividido para os terminais; a largura de banda interna da linha de alta velocidade é ocupada pelo sinal de uma das linhas de baixa velocidade por unidade de tempo, então a largura de banda é dedicada ao tráfego da próxima linha de baixa velocidade e assim por diante. O multiplexador examina os sinais das linhas de baixa velocidade com uma ordem pré-definida. A linha de alta velocidade é ocupada com um único sinal em um determinado instante. Esta técnica é assim, totalmente diferente da FDM, na qual todos os sinais são enviados ao mesmo tempo, porém cada um ocupando uma diferente porção da largura de banda.

A FDM ajusta-se naturalmente dentro do mundo dos sinais analógicos e a TDM tem sido amplamente usada com sinais digitais. A TDM é geralmente mais eficiente do que a FDM porque mais sub-canais podem ser colocados. FDM usa **guardas-de-banda** para separar frequências alocadas. Similarmente, alguma perda do tempo ocorre com TDM. As fatias de tempo alocadas devem ser separadas por um espaço de tempo umas das outras.

A TDM pode ser utilizada quando a capacidade (em quantidade de *bits* por segundo) do meio de transmissão, em muitos casos excede a taxa máxima de geração de *bits* da estações conectadas ao meio físico. Quando isto ocorre vários sinais podem ser transportados por um único caminho físico, intercalando-se porções de cada sinal durante o tempo. A TDM pode ser **síncrona** ou **assíncrona**. Na TDM síncrona o domínio do tempo é dividido em intervalos de tamanho fixo denominados *frames*, que por sua vez são subdivididos em subintervalos denominados *slots*. Cada estação deve esperar o seu *slot* dentro de cada *frame* quando então poderá transmitir dentro do tempo daquele *slot* utilizando a taxa de transmissão máxima suportada pelo meio físico. A TDM assíncrona procura eliminar o desperdício de capacidade existente no TDM síncrono, eliminando a alocação prévia do canal. Neste caso parcelas de tempo são alocadas dinamicamente de acordo com a demanda das estações, em compensação as unidades de informação devem conter um cabeçalho com os endereços origem e destino.

O único objetivo dos multiplexadores é a diminuição do custo das linhas em uma rede. Uma tendência evidente é o crescente uso de inteligência nos multiplexadores para diminuir, ainda mais o custo das linhas (aumentar a economia). Multiplexadores inteligentes podem ser vistos como equipamentos de processamento da comunicação mais gerais e incluem os concentradores e processadores *front-end*.

Um **concentrador**, em sua forma mais simples, nada mais é que um multiplexador inteligente, com memória de armazenamento e um processador. Nos multiplexadores existe uma correlação direta entre as velocidades de entrada e saída, enquanto que nos concentradores, que também incluem um software de controle, um grande número de linhas de baixa velocidade podem compartilhar um pequeno número de canais de alta velocidade. Através de armazenamento, enfileiramento e alocação estatística de canais disponíveis, os dados da fonte são "concentrados" em um canal de alta velocidade. Uma diferença fundamental entre multiplexadores e concentradores é que o primeiro trabalha com mesma taxa de entrada e saída de dados, enquanto que um concentrador inteligente pode ajustar dinamicamente suas alocações de largura de banda para modificações no tráfego.

A técnica TDM se caracteriza por alocar uma fatia de tempo  $n$  à  $n$ -ésima linha de baixa velocidade, mesmo que não exista nenhum tráfego sobre esta linha. Uma modificação para otimizar esta técnica é a **TDM estatística** (STDM) que é uma forma de concentração. Uma fatia de tempo só será alocada se existir tráfego na linha. Em geral, o uso da STDM é mais vantajoso que a TDM pois evita a má utilização da linha. No entanto, se o tráfego de várias fontes (ou destinos) é uniformemente distribuído no tempo, então o uso da TDM será menos caro e proporcionará uma melhor utilização da linha.

Um **processador front-end** é um concentrador de propósito especial que é instalado na frente do computador hospedeiro (HOST) e executa a tarefa de gerenciar a interface entre o HOST e a rede de comunicação. O principal objetivo da utilização de um processador *front-end* é aliviar o HOST das tarefas de processamento de comunicação reduzindo o tempo e o espaço dedicados pelo HOST para estas tarefas e, portanto, aumentar a sua disponibilidade para o processamento de dados.

## 5.6. Modulação

Computadores são equipamentos que armazenam, processam e codificam informações em *bits*, que correspondem a dois níveis discretos de tensão ou corrente, representado pelos valores lógicos 0 ou 1. Este tipo de informação é chamada de **digital**. As informações geradas por fontes sonoras apresentam variações contínuas de amplitude, constituindo-se no tipo de informação que comumente denominamos de **analógica**. Desta forma podemos afirmar que existem dois tipos de sinais gerados para transmissão: **sinais analógicos** e **sinais digitais**. É importante notar que qualquer tipo de informação, analógica ou digital, pode ser transmitida através de um sinal analógico ou digital. A transmissão de informação digital através de sinais analógicos é possível através de técnicas de **modulação**, que transformam os sinais digitais em sinais analógicos através da variação contínua de um de seus parâmetros.

Neste ponto, é conveniente fazer uma diferenciação entre as unidades **bits por segundo (bps)** e **bauds**. A taxa em bauds indica o número de vezes que a característica do sinal portador da informação (sinal modulado) se altera por segundo. Se o estado do sinal representa a presença ou ausência de um *bit*, então a taxa em bauds é a mesma que a taxa em bps. Por outro lado, quando usamos transmissão multinível, cada estado do sinal representa mais de um *bit* (2 *bits* na modulação *dibit* e 3 *bits* na modulação *tribit*, etc) e observamos que a taxa em bauds é menor que a taxa em bps. Por exemplo, considerando uma modulação *dibit*, onde o sinal pode estar em quatro estados e cada estado representa um *dibit* (2 *bits*), se a velocidade de sinalização for de 200 bauds, isto é, se a condição do sinal se alterar 200 vezes por segundo, teremos 400 bits transmitidos por segundo, ou seja, para cada alteração de condição do sinal, estaremos transmitindo dois bits representativos de um estado.

Um exemplo sobre a diferença entre as velocidades bauds e bps é a utilização do canal telefônico para transmissão de sinal de dados. A largura de banda deste canal é de 3100 Hz (ciclos/segundo) e na prática é usado para transmitir sinal de dados até 2400 bauds. Se desejarmos transmitir a uma velocidade de transmissão de 4800 bps neste canal, deveremos usar um sinal *dibit*, ao qual corresponderá a mesma velocidade de sinalização de 2400 bauds. Neste caso, a velocidade de transmissão é duas vezes a velocidade de

sinalização. Da mesma forma que se desejarmos transmitir 7200 bps, deveremos usar um sinal *tribit* e teremos velocidade de transmissão igual a três vezes a velocidade de sinalização.

Modulação é o processo pelo qual se imprime uma informação em uma onda portadora, pela variação de um de seus parâmetros (amplitude, frequência ou fase). O processo inverso, no qual a informação é retirada da onda portadora, é denominado **demodulação**. O equipamento que realiza a adequação dos sinais binários ao canal de transmissão é denominado **modem** (**modulador-demodulador**).

Existem três técnicas básicas de modulação:

- modulação por amplitude (AM)
- modulação por frequência (FM)
- modulação por fase (PM)

No caso específico do sinal modulador ser um sinal digital, essas técnicas tomam as seguintes denominações:

- modulação por chaveamento da fase (ASK)
- modulação por chaveamento da frequência (FSK)
- modulação por chaveamento de fase (PSK)

### 5.7. Sistemas em Banda Larga e em Banda Básica

Duas técnicas de transmissão são as mais empregadas: **sinalização em banda básica** (*baseband*) e **sinalização em banda larga** (*broadband*). Na sinalização em banda básica (ou sinalização digital) o sinal é simplesmente colocado na rede sem se usar qualquer tipo de modulação, aparecendo diretamente na rede e não como deslocamentos de frequência, fase ou amplitude de uma portadora. Sinalização em banda básica vem sendo adotada nos projetos de redes por não necessitar de *modems* e por possibilitar a transmissão em alta velocidade. Esta técnica não é adequada para transmissão a longas distâncias e para circuitos sujeitos a ruídos, interferências e erros aleatórios. Na sinalização em banda básica todo o espectro (banda) de frequência do meio é utilizado para produzir o sinal. No ambiente de redes locais a sinalização em banda básica é bastante adequada.

A sinalização em banda larga (ou sinalização analógica) realiza a multiplexação em frequência (FDM). Com esta sinalização o espectro de frequência do meio é dividido em vários canais, cada um podendo suportar diferentes tráfegos. Um caso especial de sistema de banda larga é o sistema de baixo custo que utiliza apenas um canal, denominado banda larga de único canal. A sinalização em banda larga é geralmente utilizada em redes locais. Sinalização em banda larga é geralmente utilizada em redes locais em barra. Nas redes em banda larga cada canal pode ser utilizado para transportar qualquer tipo de sinal, por exemplo, sinais analógicos, como imagem ou voz, ou sinais digitais. Os canais podem ser classificados como dedicados ou chaveados. Tanto os canais dedicados quanto os canais chaveados podem ser ponto-a-ponto ou multiponto.