APOSTILA DE FUNDAMENTOS E PRÁTICAS EM REDES DE COMPUTADORES

"Compêndio de livros, sites, monografias e contribuições pessoais"

Professor: Márcio Luiz Machado Nogueira
Versão 1.0

Índice

Apostila de Fundamentos e Práticas em Redes de Computadores

1 OBJETIVOS

Capacitar e nivelar os profissionais da área de redes de computadores.

2 COMPETÊNCIAS

C1 Fundamentos de Redes de Computadores.

C2 Fundamentos de Comunicação Digital.

C3 Modelos Referenciais em Redes de Computadores.

C4 Família de Protocolos TCP/IP

C5 Cabeamento em Redes Ethernet.

C6 Sistemas Operacionais de Redes.

3 HABILIDADES

H1 Planejar, auditar e avaliar arquitetura de redes de computadores LAN, MAN e WAN.

H2 Planejar e calcular endereçamentos de hosts e redes.

H3 Conhecer as principais tecnologias de comunicação de dados digitais.

H4 Confeccionar, reparar e avaliar cabeamentos para redes Ethernets.

H5 Instalar e configurar sistemas operacionais proprietários para redes.

H6 Instalar, configurar e auditar ativos de redes.

4 PROPOSTA DE EMENTA PARA DISCIPLINA

LAN, MAN e WAN; Parâmetros de comparações entre redes; topologias de redes; informação e sinal; transmissão analógica e digital; teorema de nyquist; lei de shannon; multiplexação e modulação; sistemas de banda larga e banda básica; comutação; técnicas de detecção de erros; meios de transmissão; ligações ao meio; arquitetura de redes de computadores; o modelo OSI da ISO; a família de protocolos TCP/IP; RS-232; EIA/TIA-568; CSMA/CD; IEEE 802.3; IEEE 802.4; IEEE 802.11; endereçamento de rede; roteamento; DNS; FTP; Telnet; HTTP; SSH; SSL; cabeamento de rede; sistemas operacionais de redes; configuração do TCP/IP; ativos de redes.

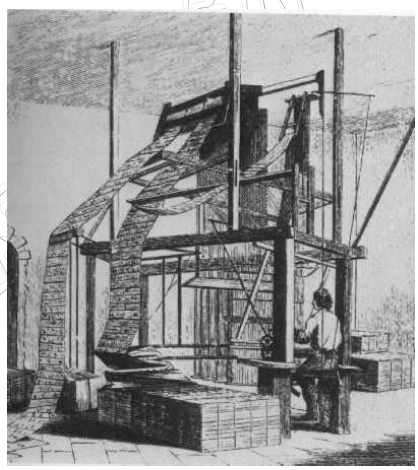
5 CARGA HORÁRIA PROPOSTA

Esta obra possui **60 horas** aula, e sendo recomendada para práticas em laboratório.

6 COMPETÊNCIA 1 – FUNDAMENTOS DE REDES DE COMPUTADORES

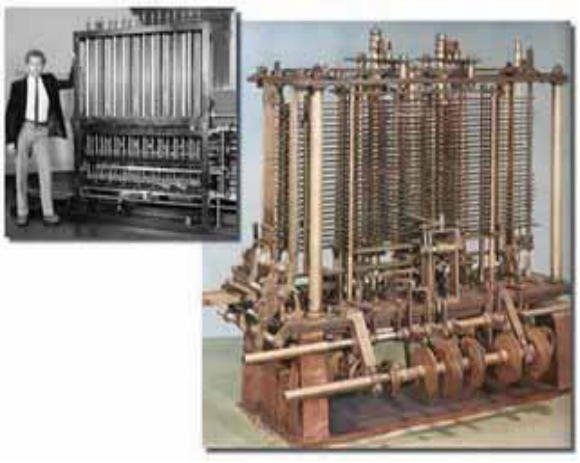
6.1 EVOLUÇÃO DOS SISTEMAS DE COMPUTAÇÃO

Para entendermos as Redes de Computadores, é necessário que observemos como se deu a evolução dos sistemas de computação, pois a procura inicial pela distribuição do poder computacional se estende até os dias de hoje.



Em 1801, na França, durante a Revolução Industrial, Joseph Marie Jacquard, mecânico frânces, inventou um tear mecânico controlado por grandes cartões perfurados. Sua máquina era capaz de produzir tecidos com desenhos bonitos e intrincados. Foi tamanho o sucesso que Jacquard foi quase morto quando levou o tear para Lyon, pois as pessoas tinham medo de perder o emprego. Em sete anos, já havia 11 mil teares desse tipo operando na França.

A origem da idéia de programar uma máquina vem da necessidade de que as máquinas de tecer produzissem padrões de cores diferentes. A idéia de Jacquard atravessou o Canal da Mancha, onde inspirou Charles Babbage (1792-1871), um professor de matemática de Cambridge, a desenvolver uma máquina de “tecer números”, uma máquina de calcular onde a forma de calcular pudesse ser controlada por cartões. O "Calculador Analítico", ou também “Engenho Analítico”, concedeu a este brilhante matemático inglês, Charles Babbage, o termo de "Pai do Computador", que inspirou a concepção de um computador atual.



Foi com Charles Babbage que o computador moderno começou a ganhar forma, através de seu trabalho no engenho analítico. O equipamento, apesar de nunca ter sido construído com sucesso, possuía todas as funcionalidades do computador moderno. Foi descrito originalmente em 1837, mais de um século antes que qualquer equipamento do gênero tivesse sido construído com sucesso. O grande diferencial do sistema de Babbage era o fato que seu dispositivo foi projetado para ser programável, item imprescindível para qualquer computador moderno.

Tudo começou com a tentativa de desenvolver uma máquina capaz de calcular polinômios por meio de diferenças, o calculador diferencial. Enquanto projetava seu calculador diferencial, a idéia de Jacquard fez com que Babbage imaginasse uma nova e mais complexa máquina, o calculador analítico, extremamente semelhante ao computador atual.

O projeto, totalmente mecânico, era composto de uma memória, um engenho central, engrenagens e alavancas usadas para a transferência de dados da memória para o engenho central e dispositivos para entrada e saída de dados. O calculador utilizaria cartões perfurados e seria automático.

Sua parte principal seria um conjunto de rodas dentadas, o moinho, formando uma máquina de somar com precisão de cinquenta dígitos. As instruções seriam lidas de cartões perfurados. Os cartões seriam lidos em um dispositivo de entrada e armazenados, para futuras referências, em um banco de mil registradores. Cada um dos registradores seria capaz de armazenar um número de cinquenta dígitos, que poderiam ser colocados lá por meio de cartões a partir do resultado de um dos cálculos do moinho.

Por algum tempo, o governo britânico financiou Babbage para construir a sua invenção. Além disso tudo, Babbage imaginou a primeira máquina de impressão, que imprimiria os resultados dos cálculos, contidos nos registradores. Babbage conseguiu, durante algum tempo, fundos para sua pesquisa, porém não conseguiu completar sua máquina no tempo prometido e não recebeu mais dinheiro. Hoje, partes de sua máquina podem ser vistas no Museu Britânico, que também construiu uma versão completa, utilizando as técnicas disponíveis na época.

Durante sua colaboração, a matemática Ada Lovelace publicou os primeiros programas de computador em uma série de notas para o engenho analítico. Por isso, Lovelace é popularmente considerada como a primeira programadora.Em parceria com Charles Babbage, Ada Augusta (1815-1852) ou Lady Lovelace, filha do poeta Lord Byron, era matemática amadora entusiasta. Ela se tornou a pioneira da lógica de programação, escrevendo séries de instruções para o calculador analítico. Ada inventou o conceito de subrotina, descobriu o valor das repetições - os laços (*loops*) e iniciou o desenvolvimento do desvio condicional.Junto com Babbage, trabalhou a jovem Ada Augusta, filha do poeta Lord Byron, conhecida como Lady Lovelace e Ada Lovelace. Ada foi a primeira programadora da história, projetando e explicando, a pedido de Babbage, programas para a máquina inexistente. Ada inventou os conceitos de subrotina, uma seqüência de instruções que pode ser usada várias vezes, loop, uma instrução que permite a repetição de uma seqüência de cartões, e do salto condicional, que permite saltar algum cartão caso uma condição seja satisfeita.

Babbage teve muitas dificuldades com a tecnologia da época, que era inadequada para se construir componentes mecânicos com a precisão necessária. Com a suspensão do financiamento por parte do governo britânico, Babbage e Ada utilizaram a fortuna da família Byron até a falência, sem que pudessem concluir o projeto, e assim o calculador analítico nunca foi construído.

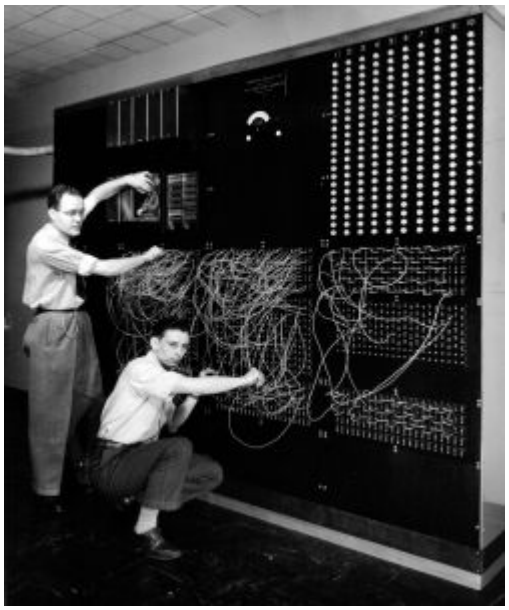
Ada Lovelace e Charles Babbage estavam avançados demais para o seu tempo, tanto que até a década de 1940, nada se inventou parecido com seu computador analítico. Até essa época foram construídas muitas máquinas mecânicas de somar destinadas a controlar negócios (principalmente caixas registradoras) e algumas máquinas inspiradas na calculadora diferencial de Babbage, para realizar cálculos de engenharia (que não alcançaram grande sucesso).

Na década de 40, em pleno auge da 2ª Guerra Mundial, o Eixo, em especial a Alemanha, passou a utilizar uma máquina eletro-mecânica, conhecida por Enigma, para criptografar suas comunicações de rádio e telégrafo. A cifra produzida por esta máquina era notavelmente consistente, concedendo a máquina o título de "Máquina Indecifrável".

Para vencer essa máquina os governos Aliados gastaram muitos esforços e dinheiro. A resposta dos Aliados vem na criação de um novo tipo de máquina, um computador analógico.

Alan Turing, professor de criptoanálise da Princeton University, e pesquisador da Bell Labs, desenvolve o Bomba, uma máquina eletromecânica de calcular, capaz de auxiliar na quebra da máquina alemã Enigma, e precursora dos primeiros computadores.

Alan Mathison Turing foi recrutado para a Escola de Códigos e Criptogramas do governo em Bletchley Park, Buckinghamshire, onde uma equipe liderada por Tom Flowers, tinha sido incumbida de decifrar os códigos militares nazistas, um trabalho urgente e secreto. Entre 1942 e 1943, Alan Turing foi enviado à Moore School e à Bell Telephone em missão secreta. Ele aperfeiçoou um sistema de codificação vocal para as comunicações telefônicas entre Roosevelt e Churchil na Bell Telephone.



Provavelmente em Princeton, Turing conheceu John von Neumann e, então, participou do projeto do ENIAC, o primeiro computador comercial da história, na universidade da Pensilvânia. Turing e seus colegas construíram o "Colossus", em Dollis Hill, ao norte de Londres, que é considerado um precursor dos computadores digitais, era enorme e, ao invés de relés eletromecânicos, usava 1500 válvulas eletrônicas, chegando a processar cerca de 5.000 caracteres por segundo. Colossus foi utilizado secretamente durante a 2ª Guerra Mundial para decifrar as comunicações do Eixo, e permaneceu secreto durante quase 40 anos. Durante esse período Alan Turing apresentou o ENIAC, motivo esse que levou muitos historiadores a acreditarem que o ENIAC havia sido o primeiro computador da história.

Com a II Guerra Mundial, as pesquisas aumentaram nessa área. Nos Estados Unidos, a Marinha, em conjunto com a Universidade de Harvard e a IBM, construiu em 1944 o Mark I, um gigante eletromagnético. Num certo sentido, essa máquina era a realização do projeto de Babbage. Mark I ocupava 120 m^2, tinha milhares de relés e fazia muito barulho. Uma multiplicação de números de 10 dígitos levava 3 segundos para ser efetuada.

Simultaneamente, e em segredo, o Exército Americano desenvolvia um projeto semelhante, chefiado pelos engenheiros J. Presper Eckert e John Mauchy, cujo resultado foi o primeiro computador a válvulas, o Eletronic Numeric Integrator And Calculator (ENIAC), capaz de fazer quinhentas multiplicações por segundo. Tendo sido projetado para calcular trajetórias balísticas, o ENIAC foi mantido em segredo pelo governo americano até o final da guerra, porém só ficou pronto em 1946, vários meses após o final da guerra. Os custos para a manutenção e conservação do ENIAC eram proibitivos, pois dezenas a centenas de válvulas queimavam a cada hora e o calor gerado por elas necessitava ser controlado por um complexo sistema de refrigeração, além dos gastos elevadíssimos de energia elétrica.

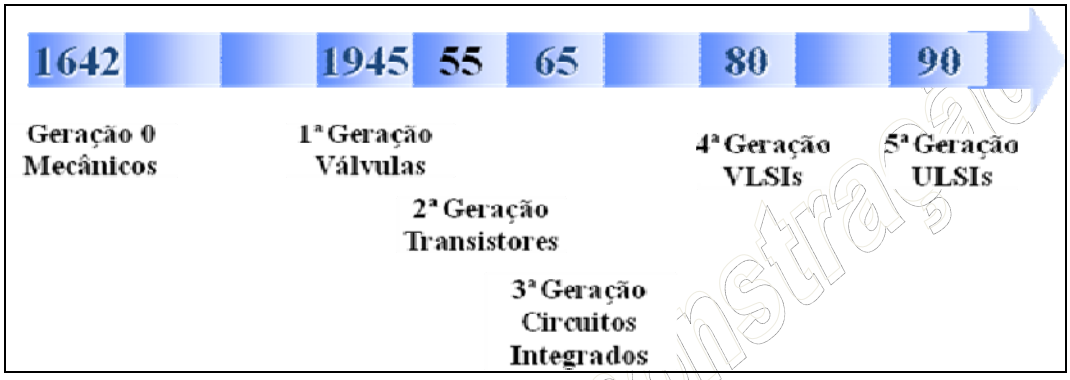
No ENIAC, o programa era feito rearranjando a fiação em um painel. Nesse ponto John von Neumann propôs a idéia que transformou os calculadores eletrônicos em “cérebros eletrônicos”: modelar a arquitetura do computador segundo o sistema nervoso central. Para isso, eles teriam que ter três características:
1 - Codificar as instruções de uma forma possível de ser armazenada na memória do computador. Von Neumann sugeriu que fossem usados uns e zeros. 2 - Armazenar as instruções na memória, bem como toda e qualquer informação necessária a execução da tarefa, e 3 - Quando processar o programa, buscar as instruções diretamente na memória, ao invés de lerem um novo cartão perfurado a cada passo.

Este é o conceito de programa armazenado, cujas principais vantagens são: rapidez, versatilidade e automodificação. Assim, o computador programável que conhecemos hoje, onde o programa e os dados estão armazenados na memória ficou conhecido como Arquitetura de von Neumann.
Para divulgar essa idéia, von Neumann publicou sozinho um artigo. Eckert e Mauchy não ficaram muito contentes com isso, pois teriam discutido muitas vezes com ele. O projeto ENIAC acabou se dissolvendo em uma chuva de processos, mas já estava criado o computador moderno.

Do surgimento de Colossus e ENIAC diversos outros computadores foram projetados. Cada novo projeto consistia em maior poder de processamento, menor consumo de energia elétrica, menor ocupação de espaço físico e menor dissipação de calor. Essa evolução ficou classificada na história da computação pelas gerações das arquiteturas dos computadores.

6.2 EVOLUÇÃO DAS ARQUITETURAS COMPUTACIONAIS

Com a divulgação dos primeiros computadores e o exponencial desenvolvimento das tecnologias pós-revolução industrial o mundo viu evoluir em 50 anos diversas máquinas magníficas. A seguir veremos as gerações dos computadores.



6.2.1 Geração 0 - Mecânicos

Até a década de 40 os primeiros vestígios de computadores traduziam-se em máquinas exclusivamente mecânicas. Caracterizadas por uma grande rigidez em termos de programas, sendo a maioria das máquinas incapazes de "trocar" de programas.

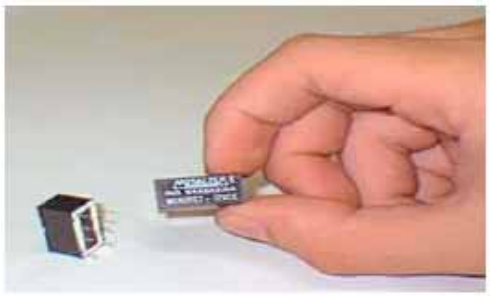
O primeiro evento que marca essa geração é a primeira calculadora mecânica do mundo, La Pascaline (a pascalina), desenvolvida em 1942 por Blaise Pascal. Sua invenção consistia em construir uma máquina capaz de realizar as quatro operações básicas da matemática, porém apenas conseguiu realizar a subtração e adição, as demais operações só podiam ser realizadas mediantes uma combinação dessas duas primeiras. O instrumento utilizava uma agulha para mover as rodas, e um mecanismo especial levava digitos de uma coluna para outra. Pascal recebeu uma patente do rei da França para que lançasse a calculadora no comércio. O engenho, apesar de útil, não obteve aceitação.

Em 1822 Babbage apresenou ao mundo seu "Engenho Analítico", já comentado na seção anterior. A Calculadora Analítica, como também era conhecida, era uma máquina de uso geral, plenamente mecânica, utilizando rodas dentadas e engrenagens, e possuia quatro componentes principais: armazenamento (memória), engenho (unidade de cálculos), leitora de cartões perfurados (mecanismo de entrada), saída perfurada e impressa (mecanismos de saída). Adotava como software da época a linguagem de montagem simples.

Em 1944, os projetos em Harvard de calculadoras automáticas através do uso de relés eletromagnéticos fizeram surgir o Mark I e Mark II, também comentados na seção anterior. Recordando que neste mesmo período existia em segredo militar o Bomba.

6.2.2 Geração 1 – Válvulas

A geração seguinte aos computadores mecânicos corresponde ao surgimento das válvulas, dispositivo que conduz a energia elétrica em um só sentido, e aprimoramento dos relés, eletroímã cuja função é abrir ou fechar contatos elétricos com o intuito de interromper ou estabelecer circuito. Em relação as máquinas mecânicas, as máquinas construídas com válvulas apresentavam maior velocidade de processamento, possibilidade de funcionamento contínuo, apresentando poucos erros de cálculo e pouco tempo de manutenção,



Os computadores dessa primeira geração caracterizavam-se por constantes quebras após muitas horas de uso, tinham dispositivos de entrada/saída primitivos, baseados quase que exclusivamente nos cartões perfurados, apresentavam uma série de desvantagens, como: custo elevado, relativa lentidão, pouca confiabilidade, grande quantidade de energia consumida e necessitavam de grandes instalações de refrigeração para dissipar o calor gerado pelo grande número de válvulas.

Os principais representantes dessa geração foram o Colossus, mantido em segredo militar, e o ENIAC, após sua publicação oficial em 1946. O ENIAC foi considerado o primeiro computador digital de uso geral, seus programas eram introduzidos por meio de cabos, que por sua vez fazia com que sua preparação para cálculos demorassem semanas. Sua arquitetura relembrava muito a do Colossus, e não por menos, seus desenvolvedores foram praticamente os mesmos. O ENIAC ocupava cerca de 170m^2, pesava mais de 30 toneladas, funcionava com 18.000 válvulas e 10.000 capacitores, além de milhares de resistores a relé, consumindo uma potência de 150Kwatts. Além de ser uma monstruosidade em tamanho, como tinha vários componentes discretos, não conseguia funcionar por muitos minutos seguidos sem que um desses componentes quebrasse, porém conseguia realizar algumas operações mil vezes mais rápido que o Mark I.

Logo em seguida ao ENIAC, Jon Von Neumann, o idealizador do conceito de programa armazenado, lança a Máquina IAS, ou mais conhecida como a Máquina de Jon Von Neumann. Um novo tipo de computador baseado em válvulas, mas que adotava a aritmética binária ao invés da decimal, e que instituiu o conceito de programa armazenado no próprio computador.

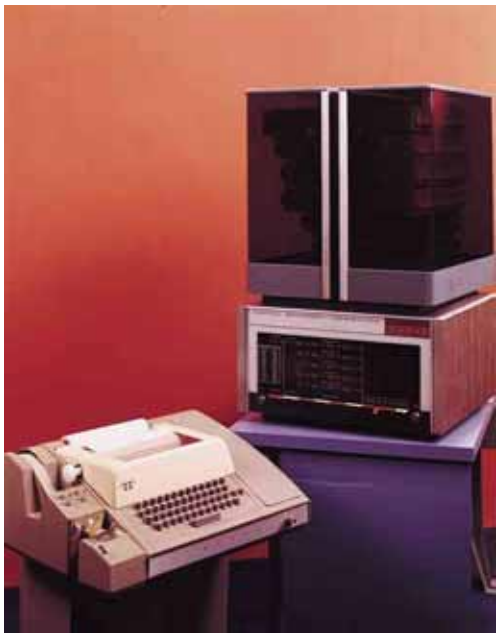
Finalizam essa época os computadores EDVAC, cuja empresa fabricante tornar-se-ia mais a frente a internacionalmente conhecida Unisys, o UNIVAC I, o primeiro computador de uso geral para fins comerciais, e o IBM 701, seguido pelos IBM-704 e IBM-709, consolidando a IBM no mercado de computadores internacionais.

6.2.3 Geração 2 - Transistores

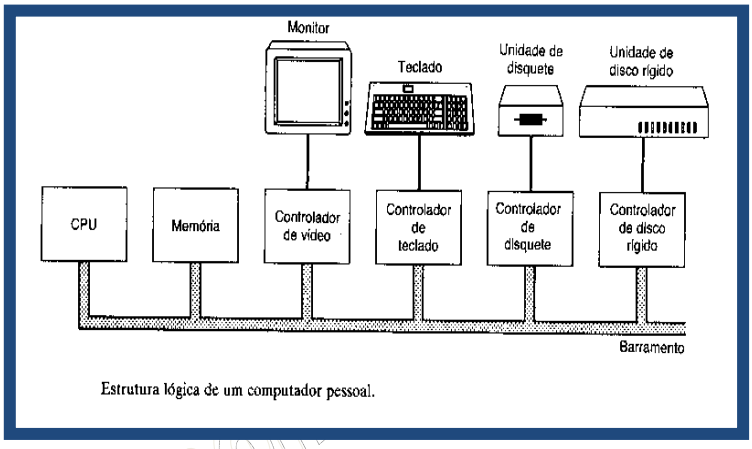
Em 1948 surgem os transistores, um amplificador de cristal usado para substituir a válvula. Apresentava um melhor custo de produção, tamanho e desempenho em relação as válvulas. Adotava uma base lógica digital equivalente a das válvulas, porém com muito menos erros de precisão. Por volta de 1957 começam a surgir os primeiros computadores experimentais a base de transistores.

Os principais representantes dessa geração são o PDP-1, da DEC, instaurando o marco da indústria de mini-computadores e de displays visuais (Monitores CRT), e os IBM 701, 7090 e 7094.

Outro marco dessa geração é a adoção das memórias com anéis ferromagnéticos, as fitas magnéticas foram a forma dominante de armazenamento secundário, permitindo uma capacidade muito maior de armazenamento e o ingresso mais rápido de dados do que as fitas perfuradas.



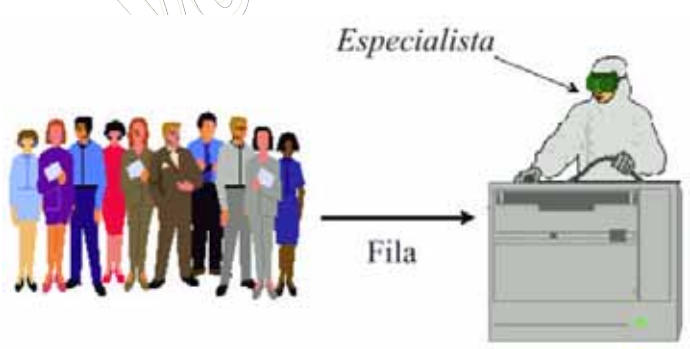
Ainda nessa mesma geração a DEC anuncia o PDP-8, adotando o novo conceito de barramento (omnibus), iniciando a era das estações de trabalho (Workstation). Abaixo a estrutura lógica do PDP-8



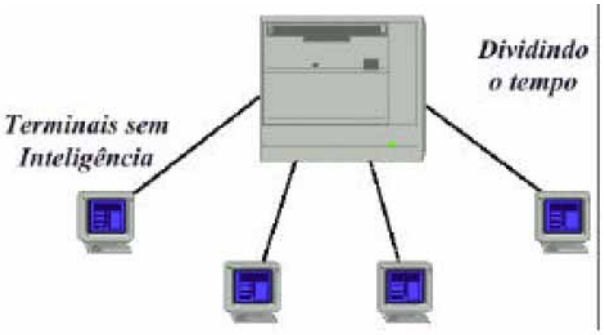
Estrutura lógica de um computador pessoal.

Outra característica dessa geração é a introdução do sistema operacional, em especial o IBM OS/360. O SO, como também é conhecido, passou a ler os cartões de: controle, programas e dados, eliminando a necessidade do programador operar a máquina diretamente. Para desenvolver os softwares para o SO, surgiram as linguagens de baixo nível Assembly e Cobol, sendo o Assembly destinado ao desenvolvimento geral e o Cobol mais voltado para o desenvolvimento de aplicações comerciais.

O surgimento desses softwares permitiu uma mudança de paradigma geral no uso dos computadores. Os sistemas de computação que eram até então caracterizados pela computação centralizada, adotando o processamento em lote (Batch), onde os usuários enfileiravam-se em uma central para submeter suas tarefas (Jobs) a um especialista (Process),



passaram a ser caracterizados pela computação centralizada, adotando o processamento de tempo compartilhado (time-sharing), permitindo que várias tarefas de vários usuários, através do uso de terminais interativos, ocupassem simultaneamente o computador central.



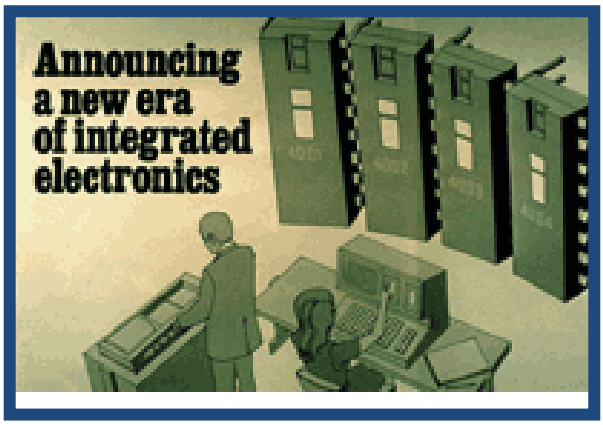
6.2.4 Geração 3 – Circuitos Integrados

Com a criação do circuito integrado, em 1961, pela Fairchild Semiconductor e Texas Instruments, as desvantagens dos computadores pareciam estar com seus dias contados. O CI, como é mais conhecido, é um circuito eletrônico constituído de elevado número de componentes arrumados em um chip de poucos centímetros ou milímetros quadrados. Permitiu a substituição de dezenas de transistores numa única peça de silício, reduzindo as dimensões dos computadores, aumentando a velocidade e reduzindo o custo de produção. Com o aumento da velocidade a unidade de tempo padrão passou a ser o nanossegundo em vez do segundo.

Essa geração também é conhecida como a evolução de tecnologia de pequena escala de integração (SSI) para média escala de integração (MSI), na qual dezenas de transistores podiam ser integrados no circuito de uma única pastilha. Também é dessa geração o surgimento dos discos magnéticos, dispositivos de armazenamentos superiores em capacidade e velocidade de acesso aos dados do que as fitas magnéticas.

Os principais representantes dessa geração são o DEC PDP-11 e o IBM 360, respectivamente:

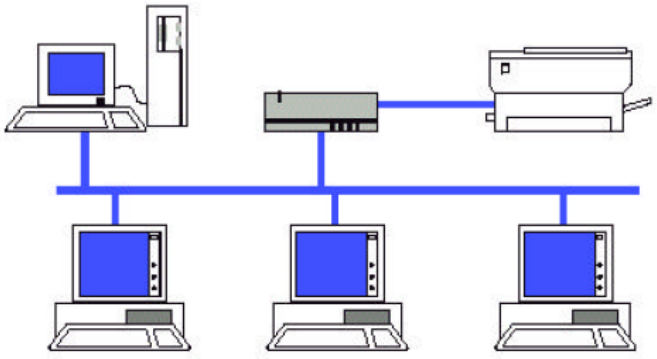
O surgimento do Intel 4004, e do Apple II, o primeiro microcomputador pessoal de sucesso comercial:



William (Bill) Gates e Paul Allen inauguram a Microsoft, comercializando os primeiros softwares para microcomputadores adaptados da linguagem de programação Basic, o M-Basic.

A IBM lança o IBM 5150, baseado no Intel 4040 e focando o mercado corporativo de estações de trabalho. Rodava o SO IBM/360 e aplicativos como M-Basic, WordStar e SuperCal.

Esta geração também marca o fim do paradigma de processamento Time-Sharing, onde a centralização das atividades (Jogs) era caracterizada pelo uso de terminais de acesso burros aos mainframes corporativos, e agora passaria a ser descentralizada. Com o aumento do poder computacional dos mini e micro-computadores, seu barateamento e as novas características de softwares que surgiam as empresas perceberam que era mais vantajoso distribuir o poder computacional ao longo da empresa, ao invés de ficar refém de um único elemento central, intolerável a falhas.



Com esse avanço, surgiu a idéia de compartilhamento de recursos, cujo objetivo é colocar os programas, equipamentos e especialmente dados ao alcance das pessoas da rede, independente da localização física do recurso e do usuário.

As redes de computadores passam então a ser motivos de estudos pelos cientistas da computação. Os elementos constituintes dessas redes eram: o serviço de rede que se desejava compartilhar, o meio físico de transmissão por onde passariam os sinais, e o protocolo, como as regras para transmissão de dados. Além disso, os seguintes componentes passaram a ser foco tanto do meio acadêmico como do mercado propriamente dito: o sistema operacional de rede, a estação de trabalho, o servidor, a placa de rede e o cabeamento.

6.2.5 Geração 4 – Escala de Integração Muito Grande (VLSI)

A quarta geração surge em 1971 com surgimento do microprocessador Intel 4004, seguido em 1972 com o Apple Macintosh, um computador voltado para o uso pessoal, cujas características inovadoras incluíam: compatibilidade, o disco ótico, processadores baseados na família Intel, sistema operacional gráfico Unix, mouse e a unidade de disquete de 3,5. A compatibilidade, que surgiu na geração anterior, é a características de poder aumentar a capacidade ou potência do hardware substituindo apenas algumas peças, e não a máquina completa.

Esta geração também é caracterizada pela migração das grandes escalas de integração - LSI (1.000 transistores por pastilha) para a muito grande escala de integração – VLSI (100.000 transistores por pastilha). Destaca-se também o surgimento da linguagem de programação C, que facilitaria a criação de novas soluções em software a partir deste momento. Todas essas mudanças possibilitaram o surgimento de diversos tipos de microcomputadores, desde notebooks até estações de trabalho mais sofisticadas:

São exemplos de destaque dessa época: o Altair 8800, considerado o primeiro microcomputador padrão mundial de uso pessoal, a série Intel de chips torna-se o padrão de mercado (8086, 8088, 80286, 80386, 80486), a IBM adota o chip Intel para o seu PC Compatible (Computador Pessoal Compatível com Hardwares Abertos) e o sistema operacional MS-DOS da Microsoft, dando início a era da microinformática (diversos modelos foram lançados, como: PC, PC-XT, PC-XT 280, PC-AT, PC-386, PC-486), o surgimento da ARPANET como a primeira rede de computadores de longa distância ou também chamada de teleinformática, disparam os usos de redes locais (LAN) e redes geograficamente distribuídas (WAN), surgem os primeiros protocolos padrões de redes, como: DECnet, SNA, TCP/IP e CCITT X.25.

Além dos desenvolvimentos voltados para o mercado pessoal, também se destacaram os desenvolvimentos voltados para a supercomputação, como: o supercomputador Cray-1, adotando processamento paralelo e máquinas vetoriais, realizava cálculos super sofisticados de manipulação de imagens, previsão de tempo, e resultados para laboratórios médicos, além de outros supercomputadores similares, como: IBM 9076 SP/2, Galaxy, Hitachi M200HIAP.

6.2.6 Geração 5 – Escala de Integração Ultra Grande (ULSI)

A quinta geração é caracterizada principalmente pela afirmação da teleinformática, a convergência entre informática e telecomunicações, além de caracterizar o momento de migração para as ultra grandes escalas de integração – ULSI (1.000.000 de transistores por pastilha).

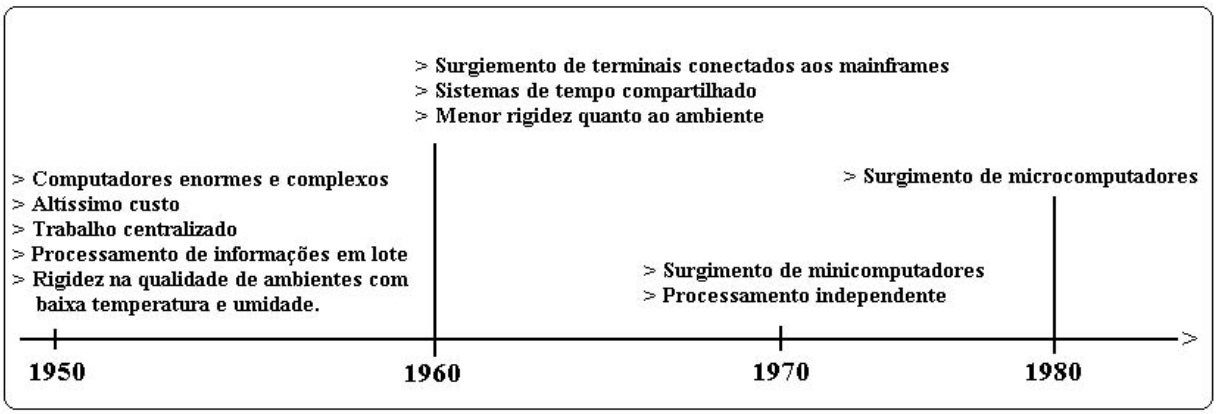
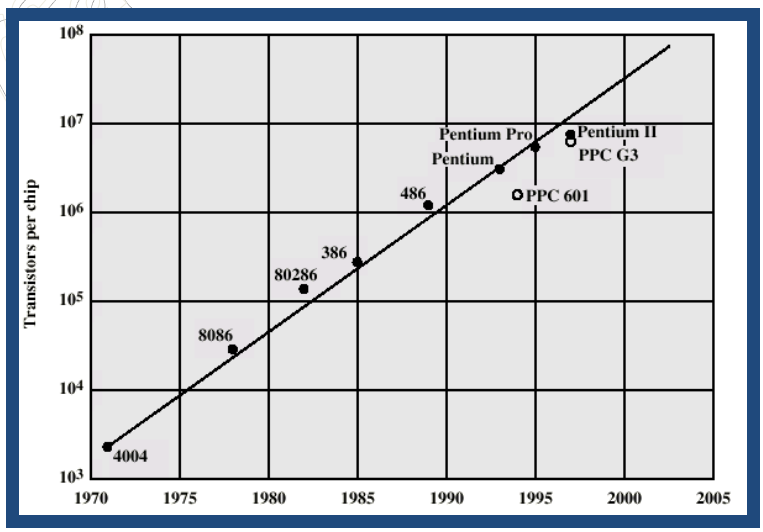


FIGURA 1: Breve evolução dos sistemas de computação

Os computadores passam a ser objetos de utilização cotidiana. A Internet se populariza, o uso de GUI (Graphical User Inteface) passa a ser uma exigência do mercado. Surgem o Intel Pentium, o processamento em paralelo dos supercomputadores passa a ser alvos de generalizações para os computadores pessoais, surgem as memórias DIMM. O pesquisador da Intel, Moore, lança sua célebre frase: a quantidade de transistores dobra a cada 18 meses, consolidando o gráfico a seguir e estimulando o crescimento das empresas de informática:

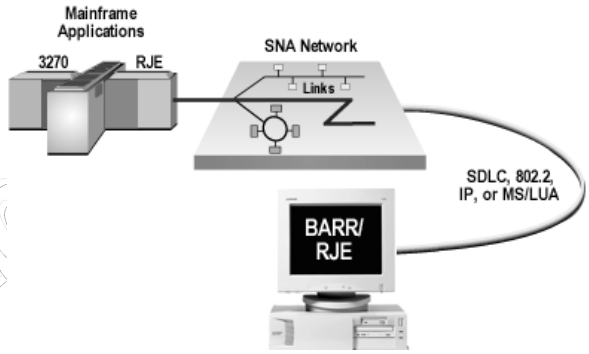


A partir de então alguns historiadores afirmam existir ainda uma sexta geração, que começaria em 2005 com a comercialização do primeiro computador quântico, entretando, em 2009, essa informação ainda não foi consolidada no meio acadêmico, sendo a priori uma tendência futura. Veremos na próxima seção os detalhamentos das redes locais e geograficamente distribuídos que surgiram deste advento.

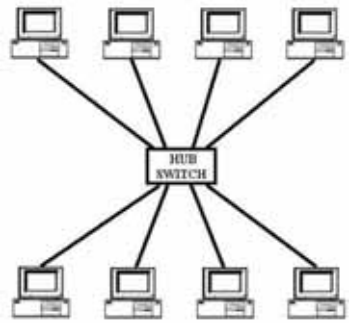
6.3 REDES DE COMPUTADORES: LAN, MAN E WAN

Vimos na seção anterior a evolução dos sistemas computacionais e onde no momento da história dos computadores surgem as redes de computadores. Nessa seção iremos aprofundar um pouco mais sobre a história, definições e conceitos das redes de computadores para nas sessões seguintes estudarmos as tecnologias que proporcionam essas redes.

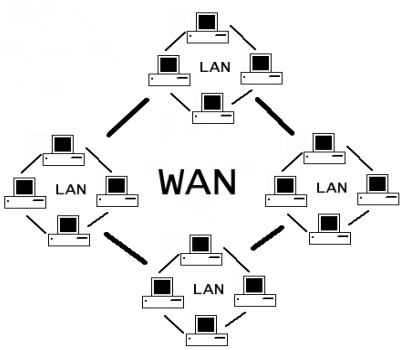
As redes de computadores surgem então por volta de 1950, na segunda geração dos sistemas computacionais. Proporcionadas pelo novo paradigma de processamento, o Time-Sharing, cuja característica era a interligação de "terminais burros" ao elo central, o mainframe. Durante esse período foram desenvolvidos padrões para o cabeamento dos terminais ao mainframe, como exemplos: cabeamento coaxial, conexões RS-232, conexões via portas LPTI. Uma característica dessas primeiras redes era a baixa velocidade, protocolos especiais foram desenvolvidos para esses fins, destacando-se o SNA (Systems Network Architecture) da IBM, este protocolo era especializado em interligar terminais remotos aos mainframes através de conexões dedicadas, surgia o conceito de áreas geograficamente distribuídas, ou WAN (Wide Area Network).



Nos anos seguintes, na terceira geração dos sistemas computacionais – a era dos circuitos integrados, as tecnologias ganharam maior poder de processamento, menor tamanho físico, e chegaram os mini-computadores e os primeiros microcomputadores. Essa nova geração permitira a migração do paradigma de processamento centralizado para o descentralizado, onde as atividades (Jobs) passariam a ser desenvolvidas distribuidamente ao longo de vários mini e microcomputadores na empresa. Com a crescente demanda de computadores nas empresas logo se percebeu a necessidade de interligá-los em rede, fortalecendo a convergência da informática com as telecomunicações.



A era da convergência, ou da teleinformática, inicia com os estudos das redes geograficamente distribuídas (WAN) e das redes locais de computadores, denominadas LAN (Local Area Network).



As primeiras LANs foram criadas no final de 1970 e eram usadas para criar links de alta velocidade entre grandes computadores centrais em um determinado local. De muitos sistemas competidores criados nessa época a Ethernet, e ARCNET eram os mais populares.

O crescimento do sistema operacioal de redes CP/M e depois dos computadores pessoais baseados em MS-DOS, proporcionaram que um único local pudesse ter dúzias e até centenas de computadores. A atração inicial das redes era geralmente compartilhar espaço em disco e impressoras à laser, os quais eram extremamente caros na época. Um entusiasmo maior com o conceito de LAN surgiu por volta de 1983, o qual foi declarado pela indústria de computadores como "o ano da LAN".

Na realidade o conceito de LAN foi estragado devido à proliferação de camadas físicas e implementações de protocolos incompatíveis, assim como confusões em como melhor compartilhar recursos. Tipicamente, cada fabricante tinha seu próprio tipo de placa de rede, cabos, protocolos e sistema operacional de rede. Uma solução apareceu com o advento do Novell NetWare, o qual proporcionou suporte a mais de 40 tipos de placas de rede e cabos, e um sistema operacional muito mais sofisticado do que qualquer um dos competidores. O NetWare dominou as LANs dos computadores pessoais até a introdução do Microsoft Windows NT Advanced Server em 1993 e o Windows for Workgroups ("Windows para grupos de trabalho").

Dos competidores do NetWare, somente Banyan Vines tinha forças técnicas comparáveis, mas Banyan nunca obteve uma base segura. A Microsoft e a 3Com trabalharam juntas para criar um sistema operacional de rede simples o qual formou a base da 3Com 3+Share Microsoft Lan Manager e IBM Lan Server. Nenhum desses particularmente teve sucesso.

No mesmo período de tempo, computadores baseados em Unix estavam utilizando redes baseadas em TCP/IP, influenciando até hoje a tecnologia dessa área.

Abaixo apresentamos um resumo cronológico dos principais eventos relacionados as redes, em especial as redes geograficamente distribuídas, que ficou mais conhecida na história como Internet:

1969 - Departamento de Defesa dos EUA contrata time de executivos, acadêmicos e pesquisadores do governo para colaborar com a ARPANET - Quatro locais escolhidos como primeiros sites da ARPANET: Universidade da Califórnia em Los Angeles (UCLA), Instituto de Pesquisas de Stanford (SRI), Universidade da Califórnia em Santa Bárbara (UCSB) e Universidade de Utah.
1970 - Criação do protocolo servidor-a-servidor NCP (Network Control Protocol), precursor do atual TCP.
1973 - Demonstração Pública da ARPANET em conferência de comunicação de computadores em Washington - Primeira conexão da ARPANET entre Inglaterra e Noruega.
1974 - Vinton Cerf e Robert Kahn definem os protocolos TCP e IP como a linguagem comum entre computadores de rede.
1975 - John Vittal desenvolve o MSG, primeiro programa de e-mail que permite encaminhar mensagens.
1976 - Mike Lesk desenvolve o Telnet que permite que duas máquinas, com sistema UNIX, se comuniquem por meio de modem e linha telefônica.
1977 - Correio eletrônico é fornecido a mais de cem pesquisadores de ciência da computação.
1978 - Vinton Cerf e Steve Crocker criam plano para separar as funções dos protocolos TCP e IP.
1979 - Especialistas da Universidade Duke estabelecem os primeiros grupos de discussão da USENET.
- Compuserve, primeiro serviço de Informação online, inicia suas operações com 1200 assinantes.

1980 - DARPA decide não tratar os protocolos TCP/IP como segredos militares e os abre a todos os interessados, gratuitamente.

1981 - ARPANET tem 213 servidores.

1982 - Departamento de Defesa dos EUA decide construir rede baseada na tecnologia da ARPANET.

1983 - ARPANET se divide em ARPANET (a Internet Comercial) e MILNET.

1984 - Sistema de domínios (DNS) é introduzido. Passa de mil o número de servidores da Internet.

1985 - símbolo ".com" é o primeiro domínio a ser registrado - Depois viriam o ".edu" e o ".gov".
- Fundada a América Online, o maior provedor Internacional de acesso à Internet até 2008.

1986 - A NSFNET cria um backbone de 56kbps (kilobits por segundo).

1987 - Número de servidores supera os 28 mil. Estabelecido o primeiro link de e-mail entre Alemanha e China. No Brasil a FAPESP (Fundação de Pesquisa do Estado de São Paulo) e o LNCC (Laboratório Nacional de Computação Científica) conectam-se a Internet dos EUA por meio de recursos próprios contratados junto a Embratel.

1988 - Estudante universitário lança o programa de vírus Internet Worm, paralisando temporariamente 6.000 dos 60 mil servidores conectados a rede.

1989 - Servidores ultrapassam a marca de 100 mil - Tim Berners-Lee começa a desenvolver o projeto World Wide Web, concluído um ano mais tarde. A WWW permite trocar informações com textos e imagens. Criação da RNP (Rede Nacional de Pesquisa), projeto voltado para coordenar e gerenciar a rede acadêmica brasileira.

1990 - Deixa de existir a ARPANET. Eletronic Frontier Foundation é criada por Mitch Kapor - World é o primeiro provedor comercial de acesso discado à Internet - Universidade de Minessota cria o navegador Gopher, que permite que internautas surfem pela rede - Cern lança a WWW.

1991 – Aprovada a implantação de um Backbone (Espinha Dorçal) para a RNP, financiada pelo CNPq. Este Backbone teria como finalidade interligar todos os centros educacionais do Brasil com a Internet americana.

1992 - Mais de 1 milhão de servidores estão conectados à Internet. Criação da Internet Society, com Vinton Cerf na presidência. Instalação do primeiro backbone brasileiro. Algumas organizações governamentais como o Ibase, também passam a ter acesso à Internet.

1993 - Marc Andreessen desenvolve o Mosaic, navegador que permite ver textos, imagens e áudio na WWW. Em um ano, mais de um milhão de cópias estavam em uso. WWW prolifera um crescimento anual de 341,6%.

1994 - ARPANET celebra seu 25º aniversário, com mais de 3 milhões de servidores conectados. Mark Andreessen e Jim Clark fundam a Netscape Communications e lançam a primeira versão do browser Netscape Navigator. Programa de buscas Yahoo! é criado por Jerry Yang e David Filo, na Universidade Stanford.

1995 - Bill Gates entra na indústria da Internet com o Microsoft Internet Explorer. Provedores de BBS com conexão discada (América Online e Prodigy) passam a oferecer acesso à Internet. Vaticano estréia site na Internet: www.vatican.va. Real Áudio permite escutar áudio em tempo real na Internet. Criação do Comitê Gestor da Internet Brasil, com o objetivo de acompanhar e coordenar o crescimento da rede no Brasil.

1996 - Cerca de 80 milhões de pessoas acessam a Internet, em aproximadamente 150 países. Número de servidores conectados chega aos 10 milhões; número de sites duplica a cada mês. A controvertida Lei da Decência nas Comunicações norte-americana proíbe a distribuição de materiais indecentes pela Internet. A

Suprema Corte declara a lei inconstitucional em 1997. Telefones via Internet chamam a atenção das empresas de telecomunicações, nos EUA, que pedem que a tecnologia seja banida.

1997 - Servidores ultrapassam a marca dos 19 milhões em todo o mundo. Microsoft lança a versão 4.0 do navegador Explorer e inclui programa em seu sistema operacional. Netscape anuncia versão 4.0 de seu navegador, o Netscape Navigator. Governo e Estados processam Microsoft por monopólio.

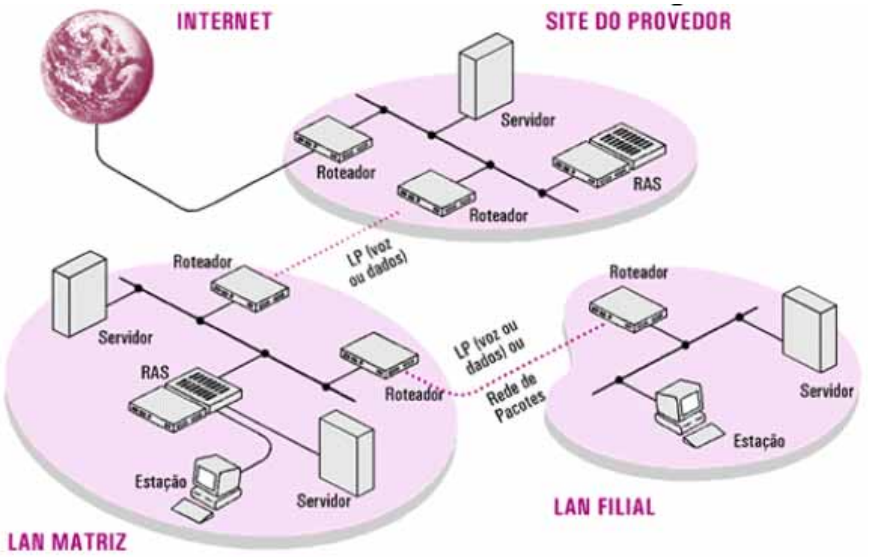
1998 - América Online chega a 12 milhões de internautas. AOL adquire Netscape.

1999 - Netscape Communicator 4.7 é lançado. Linux é sucesso como sistema operacional. Microsoft disponibiliza a versão 5.0 do navegador Explorer; é condenada por monopólio pelo governo norte americano.

Muitos outros eventos ocorreram de 1999 à 2009, porém esses vamos apresentar gradativamente ao longo do curso. Com toda essa evolução os conceitos de LAN e WAN também foram afetados, hoje, as definições mais aceitas para esses termos são:

- Rede Local (LAN – Local Area Network): são redes privadas contidas em um prédio ou em um campus universitário com alguns quilômetros de extensão. A tecnologia de transmissão quase sempre consiste em um cabo, ao quais todas as máquinas estão conectadas e apresentam uma velocidade que pode variar de 10 a 1.000 Mbps (1.000.000.000 bits), tendo um baixo retardo e cometendo pouquíssimos erros. (Ethernet, TokenRing, Token Bus);

- Rede Metropolitana (MAN – Metropolitan Area Network): é na verdade uma versão ampliada da LAN, podendo atingir muitas dezenas e poucas centenas de quilômetros com uma velocidade de até centenas de Mbps (FDDI, DQDB, ATM), apresenta como elemento principal as LP (linhas privadas) de voz ou de dados, cujas distâncias entram em conformidade com a definição das MAN;



- Redes de Longo Alcance (WAN – Wide Area Network): não apresentam limites de distância, podendo abranger uma ampla área geográfica (país ou continente). Apresentam uma taxa de erros maior do que as LAN e MAN e são normalmente de propriedade pública ou de operadoras de telecomunicações.

Assim, as redes são definidas conforme sua área de atuação. As LANs variam grandemente em tamanho – pode-se formar uma LAN a partir de dois computadores colocados um ao lado do outro na mesma sala, ou com dezenas de usuários no mesmo edifício. A parte chave na definição de uma LAN é que todos os computadores na rede estejam conectados e agrupados entre si de alguma maneira. Uma rede que se estenda por uma grande área, tal como um quarteirão, ou por um país, é conhecida como uma WAN. A figura a seguir tabela essas principais diferenças:

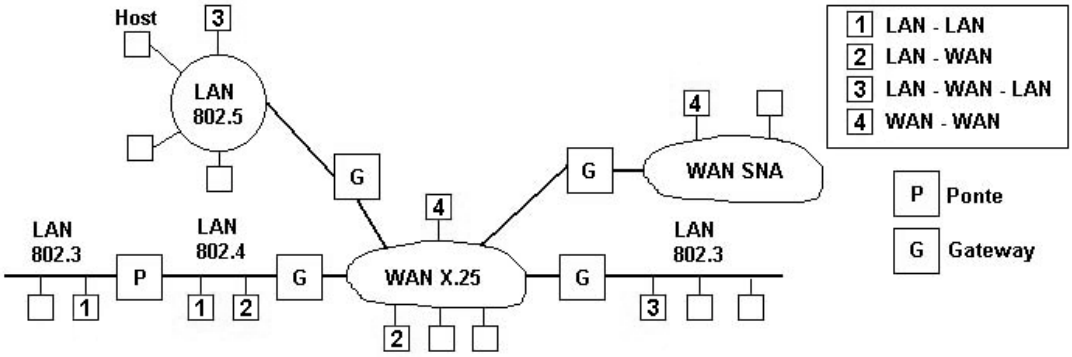
Distância entre processadores		Processadores Localizados na(o) Mesma(o)	Exemplo
0,1	m	Placa	Máquina de fluxo de dados
1	m	Sistema	Multiprocessador
10	m	Sala	Rede Local
100	m	Prédio	Rede Local
1	km	Campus	Rede Local
10	km	Cidade	Rede de Longa Distância
100	km	Estados/País	Rede de Longa Distância
1000	km	Continente	Interconexão de Redes de Longa Distância
10.000	km	Planeta	Interconexão de Redes de Longa Distância

Geralmente a arquitetura mais utilizada é a cliente/servidor. As redes deste tipo são similares em alguns pontos com as antigas redes mainframe/terminal. Em ambas, existe um computador central que é responsável pela rede e cuida de todas as solicitações. A diferença principal nos clientes de uma rede que utiliza PCs é a capacidade de processamento individual de cada estação, ao contrário dos terminais burros de uma rede mainframe/terminal cujo processamento é concentrado no mainframe (computador de grande porte – e custo elevado). Uma rede cliente/servidor é quase infinitamente expansível, ultrapassando centenas de máquinas, mesmo dezenas de milhares em uma rede de longa distância (WAN).

Outro aspecto chaves das redes é o conceito de interligação. A interligação é definida como o conjunto de arquiteturas envolvidas no processo de comunicação, podendo ser:

a) LAN-LAN: transferência de dados entre departamentos.
b) LAN-WAN: transferência de dados de uma rede a outra (ex: e-mail).
c) WAN-WAN: troca de dados entre grandes redes.
d) LAN-WAN-LAN: especialistas em locais distintos e distantes, capazes de comunicar uns aos outros.

A figura abaixo ilustra esses tipos de interligações possíveis:



Na próxima seção veremos um pouco mais sobre a história da Internet no Brasil, para compreendermos com as WANs se desenvolveram em nosso país.

6.3.1 A História da Internet no Brasil

Um dos marcos da Internet brasileira data de 1991, quando a Fapesp (Fundação de Amparo à Pesquisa de São Paulo) conseguiu estabelecer sua primeira conexão à rede mundial com protocolo IP. Mas, bem antes disso, a Internet era uma rede de pesquisa entre universidades – algo estritamente acadêmico. Em 1987, ano em que os primeiros BBSs (Buletin Board Systems – Sistema de troca de mensagens) começaram a surgir, pesquisadores e técnicos da Embratel se reuniram na USP (Universidade de São Paulo) para discutir a montagem de uma rede que interligasse universidades brasileiras e internacionais.

Não se falava em Internet, mas sim em Bitnet – uma rede de mainframes que trocava mensagens eletrônicas – e em NSFNet – rede que usava protocolos TCP/IP e que permitia, por exemplo, a transferência de arquivos (FTP). Mais tarde, ela se tornou o que conhecemos hoje como Internet.

Em 1988, o Laboratório Nacional de Computação Científica (LNCC), no Rio de Janeiro, fez a primeira conexão brasileira com uma Bitnet americana: ligou-se à Universidade de Maryland, nos EUA. Logo depois, a Fapesp se conectou ao Fermi National Laboratory (Fermilab), em Chicago.

Com o sucesso das conexões em Bitnet, surgiu a necessidade de coordenar a infra-estrutura das redes acadêmicas de computadores – interligando centros federais e esta duais. Assim, foi criada em 1989 a RNP (Rede Nacional de Pesquisa).

Depois de conseguir a primeira conexão à Internet em 1991, a Fapesp passou a ser a regulamentadora da Internet brasileira. Até hoje, ela administra os domínios (nomes para os endereços eletrônicos) e a terminação **.br**. Em 1995, passou a dividir seu poder com o Comitê Gestor da Internet do Brasil.

Com o advento da Internet comercial, inicialmente em 1986 nos EUA e 1996 no Brasil, diversos provedores comerciais de acesso à Internet surgiram (ISP – Internet Service Provider) com o objetivo de interligar hosts e redes com a Internet. Também surgiram diversos tipos de tecnologias de acesso à Internet, como a conexão discada, as linhas privadas dedicadas, as conexão à cabo, satélite, rádio, entre outras. A questão agora é como comparar o serviço de um provedor Internet (ISP) com outro? Estudaremos na próxima seção os parâmetros de comparações técnicas entre os ISP.

6.3.2 A Infra-Estrutura da Internet

Simplificadamente podemos dizer que a Internet é mantida por três elementos básicos: os provedores de backbones, os provedores locais de serviço e os usuários finais.

Um backbone (coluna dorsal) é uma rede com capacidade para transmitir grandes volumes de dados, podendo ter abrangência nacional, regional ou estadual. Para manter um backbone, um provedor deverá interligar seus computadores utilizando canais de alta velocidade, que podem ser próprios ou alugados de empresas de telecomunicações.

O backbone principal da Internet encontra-se nos EUA, sendo mantido por empresas provedoras de acesso como América Online, a Sprint e MCI. Outras empresas mantêm backbones de menor porte espalhados pelo mundo, os quais se encontram conectados ao backbone principal. Naturalmente, os donos dos backbones secundários pagam aos donos do backbone principal por estas conexões.

Um provedor local de serviço, por sua vez, paga para conectar sua rede local de computadores a um backbone e como todas as ligações entre as redes são dedicadas, forma-se uma grande rede permanente disponível.

Já os usuários finais realizarão suas conexões aos provedores através de acesso discado utilizando uma linha telefônica com um modem para acessar a Internet ou qualquer outra tecnologia de acesso à Internet.

Com relação aos custos, este é fragmentado conforme as ligações realizadas entre backbones, provedores e usuários onde o usuário final efetua o pagamento de uma parcela dos custos de manutenção dos serviços oferecidos pelo provedor. Por sua vez, o provedor realiza o pagamento da locação do link com o backbone e assim sucessivamente.

6.4 PARÂMETROS DE COMPARAÇÕES ENTRE REDES

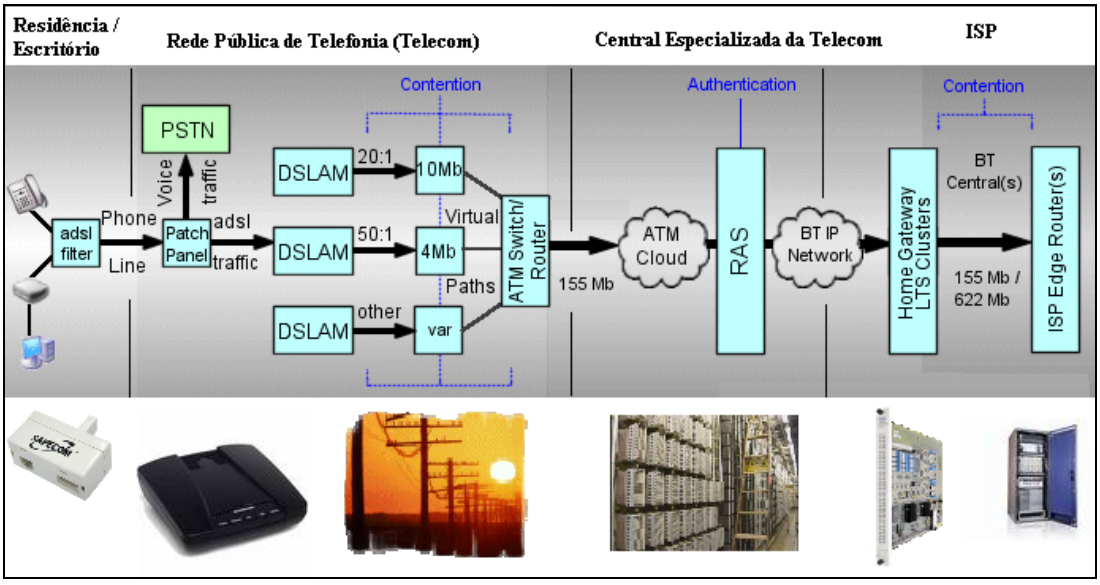
Para escolher um tipo particular de rede que suporte um dado conjunto de aplicações, nem sempre a melhor solução é a mais adequada, pois vários itens devem ser analisados, como:

- Custo
- Retardo de transferência
- Desempenho
- Confiabilidade
- Modularidade
- Compatibilidade
- Sensibilidade tecnológica

6.4.1 Custo

O custo da rede pode variar muito de acordo com o que vamos usar em termos de *hardware* e *software*, por isso devemos nos importar com a relação custo/benefício do material proposto para a implementação da rede. O custo é dividido entre o custo das estações de processamento (microcomputadores, minicomputadores, etc), o custo das interfaces com o meio de comunicação e o custo do próprio meio de comunicação. Por exemplo, uma conexão banda larga ADSL, temos: o custo do equipamento provedor do serviço no ISP (quanto mais caro for este equipamento mais oneroso ficará a prestação de serviços), o custo do meio físico por parte da operadora contratada pelo ISP (cabeamento, manutenção, monitoração), e o custo do equipamento/interface (se necessário) que o cliente precisará adquirir para acessar o meio físico ao ISP. Dessa forma, o provimento de uma tecnologia como ADSL pode ser muito mais onerosa do que uma simples conexão discada (Dial-UP), do qual não podemos aguardar uma equivalência ou equiparação de valores no mercado.

Ainda aproveitando o exemplo do ADSL vejamos um detalhamento dos custos deste tipo de tecnologia tanto para o ISP (prestador do serviço) quanto para o usuário (o cliente do serviço):



O diagrama anterior representa uma rede genérica de tecnologia ADSL. Nesta rede temos como o primeiro extremo a rede do ISP (lado direito da figura), do qual necessita de um Roteador Edge para tratar os protocolos ADSL para sua própria rede local, em seguida esse Roteador Edge é ligado a um modem de comunicação com a operadora de telecomunicações locais, o Home Gateway. A Central especializada da operadora de Telecom interligará a LAN do ISP com sua LAN-PSTN (rede pública de telefonia), distribuindo o serviço do ISP para todos os clientes da Telecom. No segundo extremo observamos o usuário (residencial, escritório ou empresarial), onde surge a presença de um Filtro ADSL, necessário para a devida separação do sinal, e o modem ADSL, como comunicação de dados sobre a rede PSTN.

Caso o serviço ofertado pelo ISP fosse apenas uma conexão Dial-UP (discada), os custos com o Roteador Edge, o Home Gateway, a Central Especializada da Telecom, o Filtro Adsl e o Modem Adsl, seriam desprezados, e com isso o valor da manutenção a ser cobrado mensalmente para o cliente seria consideravelmente reduzido. Dessa forma, o custo está diretamente relacionado as tecnologias, onde teremos tecnologias com valores mensais mais accessíveis e outras mais onerosas.

6.4.2 Retardo de transferência

É o tempo que a mensagem leva desde a sua geração pela estação de origem até chegar na estação de destino. Esse retardo pode ser decomposto em:

- **Retardo de acesso**: tempo desde que a mensagem é gerada na estação de origem até o início de sua transmissão pelos meios de comunicação, ou seja, o tempo que a estação de origem leva para conseguir a “vez” no meio de transmissão, como exemplo: aguardar até que a discagem complete a ligação;
- **Retardo de Transmissão**: tempo que a mensagem leva desde o início da transmissão pelo meio físico até a sua chegada na estação de destino, como exemplo: o “*ping”* de um pacote.

6.4.3 Tempo de Resposta

É o tempo que a mensagem leva desde a sua geração pela estação de origem, chegada na estação de destino, que retorna uma mensagem de confirmação para a estação de origem. É uma composição do retardo de transferência nos dois sentidos da comunicação somados ao tempo de processamento da mensagem nos dois lados da comunicação, geralmente mensurado em termos de TTL (Time-to-Live), sendo:

EMISSOR	DESTINATÁRIO	EMISSOR
Processamento + Retardo de Transferência + Processamento + Retardo de Transferência + Processamento		

Para exemplificar, podemos imaginar que o emissor envia uma mensagem para o destinatário, porém, a mensagem ao chegar no destinatário encontra um congestionamento no computador em função de uma aplicação multimídia que esteja sendo executada. O tempo de processamento então no destinatário irá elevar o tempo de resposta da comunicação.

6.4.4 Desempenho

O desempenho é a conseqüência da seleção de um mecanismo de interconexão orientado para a natureza da aplicação. Em outras palavras, seria a capacidade de tráfego para um determinado meio limitado. Vamos exemplificar:

Supondo uma aplicação computacional que necessita de um canal de acesso a LAN em 100Mbits. Caso a LAN tenha disponibilidade de fornecimento de 100Mbits exclusivo para essa aplicação podemos dizer que o desempenho da rede é de 100%. No entanto, a realidade é um pouco diferente, a LAN pode conter diversos grupos de estações de trabalhos, diversas ligações físicas e não conseguir garantir 100% dos 100Mbits para essa aplicação. Supondo que houve uma monitoração desse tráfego, e que ficou constatado que a vazão média da LAN para esta aplicação em especial ficou em 80Mbits, podemos dizer que o desempenho na rede foi de 80%.

Dessa forma, podemos compreender que o desempenho está diretamente associada a limitação do meio de comunicação. Desempenhos abaixo de 100% implicam em pontos de congestionamento, e desempenhos acima de 100% implicam em folgas no meio de comunicação. Exemplificando:

Dada uma LAN de capacidade 10Mbits, uma aplicação A de consumo de 1Mbit, e uma aplicação B de consumo de 50Mbits. Nessa conjectura, podemos dizer que a LAN possui desempenho de 1000% para a aplicação A, apresentando considerável folga para essa aplicação, e um desempenho de 20% para a aplicação B, apresentando considerável congestionamento para essa aplicação.

6.4.5 Confiabilidade

É calculada levando em consideração o tempo médio ocorrido entre as falhas (MTBF – Mean Time Betewen Failures) de transmissão ou processamento, e sua restauração (MTTR – Mean Time to Repear) e tempo de reconfiguração após falha (MTRF – Mean Time do Recovery Failures). Implica em uma arquitetura tolerante a falhas, contendo um número elevado de componentes idênticos, a fim de proporcionar uma ótima estrutura redundante sem o custo de aquisição de equipamentos espelhados para contingência, e pode ser compreendida como uma fórmula, como exemplo:

Confiabilidade = MTBF + MTTR + MTRF / Tolerância a Falhas

Neste exemplo de confiabilidade, que também pode assumir diversos outros vetores de configuração, implicaria dizer que uma arquitetura não tolerante a faltas (zero de tolerância) é uma utopia, visto que na fórmula não podemos ter denominador zero, dessa forma, um mínimo que seja de tolerância deve ser previsto, ainda mais que não existe uma arquitetura ou solução 100% tolerante a falhas.

Outro aspecto é que precisamos lidar com uma unidade única, neste caso o tempo. Definir uma confiabilidade em segundos, minutos, horas, dias, semanas, meses, anos, fica a critério do usuário, porém, uma vez definida a unidade de tempo da confiabilidade, todos os demais tempos deverão ser convertidos para essa mesma unidade.

Continuando, uma arquitetura supostamente com índice de tolerância de 0,01/mês (limite de 1% de falhas ao mês), quase próxima a zero, cujos tempos entre falhas, de restauração e de reconfiguração sejam também mínimos, poderia expressar uma confiabilidade em horas da seguinte forma:

Confiabilidade (horas) = 2h + 4h + 1h / (30 dias * 24 horas / 100) = **0,97222**

Com esse referencial-hipotético, podemos agora analisar alguns casos para compreendermos melhor o índice de confiabilidade.

Vamos supor que uma segunda empresa apresenta os mesmos tempos entre falhas, restauração e reconfiguração, porém, devido seu negócio, apresenta uma maior tolerância a falhas, o que implicaria dizer que falhas mais prolongadas do que no primeiro exemplo, não acarretariam prejuízos para esta empresa, vejamos os números então para uma tolerância de 5%/mês:

Confiabilidade (horas) = 2h + 4h + 1h / (30 dias * 24 horas * 5 / 100) = **0,19444**

Comparando os resultados, percebemos que a confiabilidade para a segunda empresa se afastou de 1, de fato, o número 1 seria o modelo utópico. Dessa forma, quanto mais se aproximada de 1 o índice de confiabilidade, significa que a empresa acredita mais no serviço prestado.

Vejamos um último exemplo, vamos supor agora que uma terceira empresa apresenta o mesmo critério de tolerância a falhas da primeira empresa (1%/mês), porém que apresenta um serviço de mais baixa qualidade, cujos tempos entre falhas e restauração sejam maiores, ficaria então:

Confiabilidade (horas) = 4h + 6h + 1h / (30 dias * 24 horas * 1 / 100) = **1,52777**

Percebemos então que o índice de confiabilidade para a terceira empresa é de 1,5 horas, ou seja, muito superior ao índice de 0,9 horas do primeiro exemplo. Entretanto este índice pode confundir-se com o 0,2 horas do segundo exemplo. Neste caso, precisamos compreender que o ponto de equilíbrio da fórmula está compreendido no numeral 1, ou seja, valores que se afastam desta tara implicam em conseqüências negativas. Nos casos de tolerância, a afastabilidade se dará para valores abaixo de zero, nos casos de tempo de serviço a afastabilidade se dará para valores acima de zero.

Um bom observador poderia argumentar que uma alta tolerância a falhas e um alto tempo de serviço poderia implicar em um índice agradável de confiabilidade, vejamos:

Confiabilidade (horas) = 12h + 48h + 12h / (30 dias * 24 horas * 10 / 100) = **1**

Segundo o exemplo acima, o índice de confiabilidade para a prestação de um serviço com 12 horas entre falhas, 48 horas de restauração e 24 horas de reconfiguração é completamente aceitável e perfeito para a empresa cuja tolerância a falhas seja de 10%/mês. Esse resultado é totalmente condizente com o perfil da empresa e demonstra que a fórmula apresenta flexibilidades para os diversos perfis de usuários.

6.4.6 Modularidade

Capacidade de ampliar um sistema sem afetar as aplicações existentes, apresentando facilidade na mudança de *hardware* e para acrescentar mais componentes (crescimento), também visto em termos de manutenabilidade como um facilitador, ao dividir as partes maiores em menores possibilitando a especialização de equipes distintas;

6.4.7 Compatibilidade

Conhecida também como interoperabilidade, é a capacidade do sistema de interligar-se a dispositivos de outros fabricantes quer no nível de *software* ou *hardware*.

Um exemplo seria a compatibilidade entre redes de acesso a Internet. Durante os anos de 1992 à 1995 usuários da RNP (Rede Nacional de Pesquisa) não conseguiam se comunicar com os usuários da TeleBras. Apesar de ambas as redes estarem ligadas a Internet americana, entretanto seus backbones não possuíam compatibilidades de acesso, visto que os protocolos em uso eram incompatíveis.

Outro exemplo seria a compatibilidade de uma determina interface de acesso ao meio. Muitos usuários desistem de migrar para uma nova tecnologia de acesso ao meio, como Rádio, ADSL, ISDN (veremos esses conceitos mais a frente), pelo fato do equipamento (modem) adquirido ser incompatível com as demais tecnologias.

Por fim, o ISP pode adquirir um equipamento cuja propriedade incompatibilize com o equipamento do seu concorrente. Desta forma, usuários de um ISP-A podem não se comunicar com os usuários do ISP-B em função de equipamentos de mesma tecnologia porém de incompatibilidade de interoperabilidade.

6.4.8 Fatores não técnicos

Muito importante também na escolha de uma tecnologia é conhecer aspectos não técnicos que podem comprometer a prestação do serviço. Aspectos como localidade, cultura, política, fatores naturais, podem diretamente afetar o serviço. Vejamos:

Vamos supor um ISP cujo serviço seja Internet à Cabo. Uma característica técnica deste tipo de empresa é sua distribuição de cabos pelos postes da cidade. Agora vamos supor uma localidade como uma região metropolitana, cujo índice de acidentes envolvendo carros e postes seja alto. A prestação de serviços da Internet à Cabo ficará completamente comprometido, pois o tempo médio entre falhas (MTBF) aumentará consideravelmente, afastando o índice de confiabilidade de sua tara.

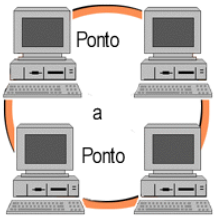
Outro exemplo poderia ser um ISP à Rádio. Uma característica técnica deste tipo de empresa é sua distribuição através da criação de pontos de repetição sobre os prédios mais altos da localidade. Agora vamos supor uma localidade como uma região do interior, cujo índice de vandalismo envolvendo tiroteios e antenas de rádio como alvo seja alto. Da mesma forma que vimos no exemplo à Cabo o MTBF deste ISP ficará comprometido.

Em fim, diversos outros parâmetros de comparações ainda existem, como: disponibilidade, facilidade de desenvolvimento, dispersão geográfica, complexidade lógica, facilidade de uso, etc. O importante é você analisar quais os aspectos técnicos e não técnicos existem no seu negócio, quais as garantias de serviços você precisa, após isso firmar um contrato denominado SLA (Service Level Agreement) com seu ISP. Muito dificilmente você encontrará um provedor de serviços que lhe assegure através de um termo SLA a qualidade de todos os itens técnicos de que você necessita. Frente a esta situação, a criação de planos de continuidade de negócio, recuperação de desastres e análises de riscos contínuos precisam ser elaborados cuidadosamente.

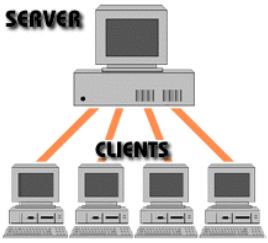
Na próxima seção apresentaremos os diferentes tipos de projetos envolvendo as LAN, MAN e WAN. Conhecendo esses tipos de projetos você conhecerá também algumas características técnicas nativas de determinadas tecnologias, facilitando sua pesquisa por parâmetros de comparações.

6.5 LINHAS DE COMUNICAÇÃO

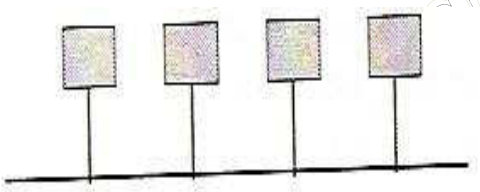
Ao organizar os enlaces físicos (interligações entre redes) num sistema de comunicação, confrontamo-nos com diversas formas possíveis de utilização das linhas de comunicação. Em primeiro lugar as ligações físicas podem ser de dois tipos: ponto a ponto ou multiponto.



Ligações ponto-a-ponto caracterizam-se pela presença de apenas dois pontos de comunicação, um em cada extremidade do enlace.

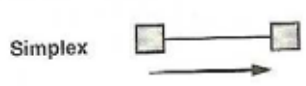


Nas ligações multipontos, que em redes são mais conhecidas como cliente-servidor, observa-se a presença de três ou mais dispositivos de comunicação com possibilidade de utilização do mesmo enlance.

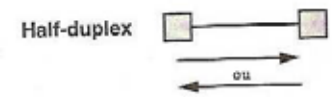


As ligações multipontos também apresentam a característica dos nós poderem ser interligados usando apenas um circuito tronco. Quando há um conjunto de troncos, por onde os nós possam se comunicar de forma contigencial ou escalonável, dizemos se tratar de um entroncamento.

A forma de utilização deste meio físico, ponto-a-ponto ou multiponto, dá origem à seguinte classificação sobre a comunicação no enlace:



Simplex: o enlace é utilizado apenas em um dos dois possíveis sentidos da transmissão;



Half-duplex: o enlace é utilizado nos dois possíveis sentidos da transmissão, porém apenas um por vez;

Full-duplex: o enlace é utilizado nos dois possíveis sentidos da transmissão simultaneamente.

O modo Simplex de comunicação geralmente é empregado para difusão, como emissoras de rádio AM/FM, televisão convencional e multicast em redes (veremos ainda mais adiante). Sua característica é a alta velocidade, visto que não apresenta recursos de controle da comunicação que geralmente atrapalham o desempenho.

O modo Half-duplex apresenta como característica principal o mecanismo de troca de sentido (*turn-around*), um controlador capaz de identificar em qual sentido a comunicação deve fluir. São exemplos desse modo de comunicação os rádios trunks, utilizados em taxis, empresas de segurança, e técnicos de campo. Em redes de computadores o modo half-duplex se apresenta tradicionalmente nos equipamentos de rede sem fio, onde os rádios são configurados ou para receber ou para enviar sinais.

O modo Full-duplex é uma combinação das melhores práticas dos modos anteriores. Ele apresenta alta velocidade, possui um *turn-around* automático e que não degride o desempenho da comunicação. Em compensação apresenta o maior custo de implementação e manutenção. Um detalhe do modo *full-duplex* é que ele pode se apresentar de forma física, um único cabeamento compartilhando o mesmo caminho em ambos os sentidos, isso é possível através da multiplexação, que veremos mais adiante, ou de forma lógica, onde na planta do projeto aparece apenas uma ligação, mas na verdade sendo pelo menos dois cabeamentos distintos, duas comunicações simplex, uma para envio e outra exclusiva para recepção.

Cada um desses modos de comunicação pode ser utilizado em cada tipo de linha de comunicação, essa combinação gera o que chamamos de topologias de linhas de comunicação, ou mais especificamente na nossa área: as topologias de redes de computadores. Estas topologias, conforme sua abrangência (LAN, MAN ou WAN), apresentarão características técnicas distintas.

6.6 TOPOLOGIAS DE REDES DE COMPUTADORES

A seção anterior apresentou os tipos e modos das linhas de comunicação, que combinados dão origem as chamadas topologias de redes de computadores. Essas topologias estão dividas conforme a área de abrangência, sendo:

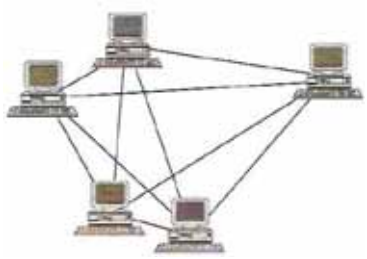
- Redes Geograficamente Distribuídas
- Redes Locais e Metropolitanas

6.6.1 Redes Geograficamente Distribuídas

Nas redes WAN encontramos os seguintes tipos de topologias:

- Topologia *Full-Mesh* – ou redes totalmente conectadas;
- Topologia em Anel;
- Topologia *Partial-Mesh* – ou redes parcialmente conectadas;
- Redes Comutadas por Pacotes;

A **topologia *Full-Mesh***, ou redes totalmente conectadas, é a primeira e mais intuitiva de todas na criação de um projeto de redes. Nessa topologia cada estação se conecta a todas as demais estações, sem exceção. Dessa forma, cada estação conhece exatamente o caminho para alcançar a outra estação, sem mediadores.

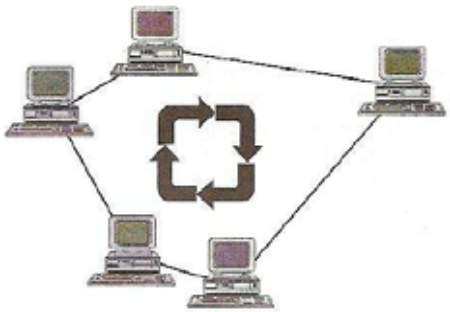


Adota uma linha de comunicação multiponto e pode operar através de qualquer modo de comunicação, mas preferencialmente a *full-duplex*.

Apesar de sua segurança, pelo fato de que a interrupção em uma estação não compromete a comunicação das demais, entretanto é a que apresenta maior custo, pois para cada nova estação na rede precisaremos incluir uma nova interface de rede nas demais estações. Logo, podemos concluir que esta topologia é ideal em redes minúsculas, porém impraticável em redes com mais de 5 nós. Por exemplo, seriam necessárias N(N-1)/2 ligações ponto-a-ponto para que se pudesse conectar todos os pares de estações através de linhas dedicadas. Dessa forma, o custo do sistema, em termos de instalação de cabos e de hardware específico para comunicação, cresceria com o quadrado do número de estações, tornando tal topologia economicamente inviável.

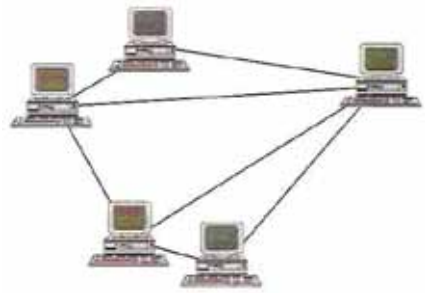
A próxima topologia, a **topologia em anel**, retrata a extremidade oposta da topologia *full-mesh*. Nessa topologia procura-se diminuir ao máximo o número de ligações no sistema além de simplificar ao máximo o tipo de ligação utilizada.

Dessa forma, utilizam-se, em geral, ligações ponto a ponto que operam num único sentido de transmissão (*simplex*) fazendo com que o anel apresente uma orientação ou sentido único de transmissão como o indicado pelas setas da ilustração. Uma mensagem deverá circular pelo anel até que chegue ao módulo de destino, sendo passada de estação em estação, obedecendo ao sentido definido pelo anel.



Apesar de representar uma economia considerável no número de ligações, em sistemas geograficamente distribuídos tal topologia apresenta fatores limitantes que inviabilizam a sua utilização. O primeiro deles diz respeito ao aumento de pontos intermediários entre os pontos finais da comunicação. Em redes geograficamente distribuídas isso significa um aumento drástico no numero de ligações pelas quais uma mensagem tem que passar até chegar ao seu destino final, ou seja, um aumento intolerável no retardo de transmissão, particularmente no caso de redes geograficamente distribuídas com meios de transmissão de baixa velocidade. Outro fator limitante refere-se à inexistência de caminhos alternativos para o tráfego das mensagens, em redes geograficamente distribuídas caminhos alternativos devem ser providenciados, principalmente se as linhas utilizadas forem de baixa velocidade e pouca confiabilidade, o que é o caso da maioria das redes existentes.

Considerando as limitações de velocidade e confiabilidade somos levados, naturalmente, à

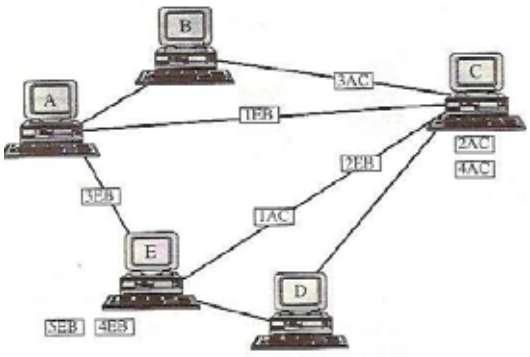


introdução de caminhos redundantes para um aumento tanto de confiabilidade quanto de desempenho através do paralelismo de comunicações, sem, no entanto, cairmos na topologia completamente ligada, que possui as restrições antes apresentadas. Somos levados, assim, a uma topologia intermediária, que é utilizada pela maioria das redes geograficamente distribuídas: **a topologia parcialmente ligada**, também conhecida como topologia em grafo.

Nessa topologia, nem todas as ligações entre pares de estações estão presentes, mas caminhos alternativos existem e podem ser utilizadas em casos de falhas ou congestionamentos em determinadas rotas. No caso em que estações sem conexão física direta desejem se comunicar, elas deverão, de alguma forma, encaminhar as suas mensagens para alguma outra estação que possa fazer a entrega da mensagem para a estação de destino. Esse processo pode se repetir várias vezes, de forma que uma mensagem pode passar por vários sistemas intermediários até chegar ao seu destino final.

A comunicação entre dois módulos processadores (chamados *Equipamentos Terminais de Dados* – ETDs ou *Data Terminal Equipments* – DTEs) pode ser realizada por chaveamento de circuitos, chaveamento de mensagens ou chaveamento de pacotes (como veremos com mais detalhes adiante). Em sistemas por chaveamento (ou comutação) de circuitos, um canal entre o ETD fonte e o ETD de destino é estabelecido para uso exclusivo dessas estações até que a conexão seja desfeita, de maneira idêntica a uma chamada telefônica. Chaveamento de mensagem ou de pacote vai otimizar o uso dos meios de comunicação, tentando evitar a monopolização de todo o caminho durante uma conversação.

Em sistemas por chaveamento de mensagem, a mensagem por completo é enviada ao longo de uma rota do ETD fonte ao ETD de destino. Em cada nó do caminho, a mensagem é primeiro armazenada, e depois passada à frente, ao próximo nó, quando o canal de transmissão que liga esses nós estiver disponível. Sistemas por chaveamento de pacote diferem dos de chaveamento de mensagem pelo fato da mensagem ser quebrada em quadros ou pacotes antes da transmissão ser efetuada.



A transmissão de cada pacote pode ser feita por um único caminho ou por caminhos diferentes, sendo a mensagem reagrupada quando chega ao destino, conforme pode ser visto na imagem ao lado.

Tanto na comutação de pacotes quanto na comutação de mensagens não existe a alocação de um canal dedicado da estação fonte à de destino, de uso exclusivo da comunicação, como no caso da comutação de circuitos.

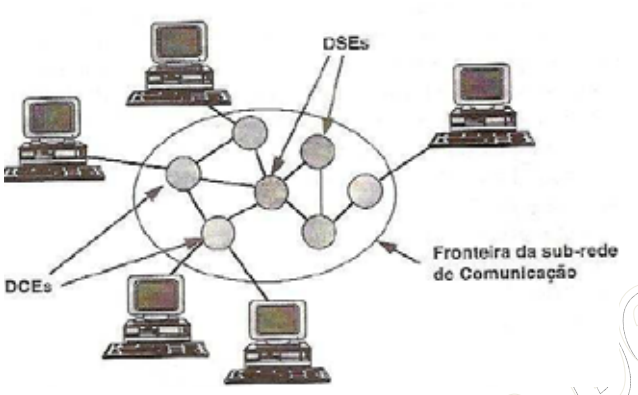
A escolha do caminho fim a fim, isto é, do módulo (nó da rede) de origem ao nó de destino, por onde uma mensagem deve transitar (tanto na comutação de circuito, quanto na de mensagem ou de pacotes), é comumente chamada de roteamento. A escolha da rota pode ser feita a priori, antes do envio da mensagem, ou ser realizada passo a passo. No primeiro caso, diz-se que é estabelecida uma conexão entre os nós de origem e destino e, neste estabelecimento, é definida a rota por onde deverão transitar as mensagens enquanto perdurar a conexão. No segundo caso, pode haver ou não o estabelecimento de conexão mas, independente disso, cada nó intermediária do caminho fim a fim é responsável pela escolha do próximo nó do caminho no instante que recebe a mensagem a despachar, e não a priori, como no caso anterior.

Vários algoritmos de roteamento já foram propostos e são, na sua maioria, baseados na manutenção de tabelas em cada um dos MPs. Voltaremos a falar de roteamento mais a frente. Muitas das características desejáveis de uma comutação resultam do uso de roteamento adaptável. Nesse roteamento, o caminho de transmissão entre dois pontos da rede não é preestabelecido, mas escolhido dinamicamente, com base nas condições da rede no tempo de transmissão. Com essa capacidade de alocação de recursos (rotas) baseada nas condições correntes, a rede é capaz de contornar efeitos adversos tais como um canal ou dispositivo de comunicação sobrecarregado, ou ainda, uma falha de componentes.

Todos os módulos processadores (ou estações) devem ser capazes de reconhecer se uma mensagem ou pacote a eles entregue deve ser passado para outra estação, ou se tem como destino a própria estação. Qualquer rede com topologia diferente da totalmente ligada tem a necessidade de definir mecanismos de endereçamento que permitam aos MPs decidir que atitude deve tomar ao receber uma mensagem ou pacote. Esse endereçamento irá consistir em uma forma de identificar univocamente cada uma das estações conectadas à rede. No caso de ser estabelecida uma conexão entre dois nós da rede antes da troca de qualquer mensagem, o endereço dos nós de origem e destino só são necessários quando do estabelecimento da conexão. A partir daí, basta que as mensagens ou pacotes transmitidos carreguem consigo a identificação da conexão para que o encaminhamento seja feito a contento. Por outro lado, caso não haja estabelecimento de conexão, cada pacote ou mensagem deve carregar o endereço do nó de destino e de origem.

Em redes por chaveamento de pacotes, varias tarefas devem ser realizadas por uma estação. Uma delas é a escolha do caminho que deve seguir cada pacote, ao que demos o nome de roteamento; outra é o armazenamento dos pacotes recebidos de outras estações, que devem prosseguir seu caminho, e dos seus próprios pacotes a serem transmitidos; outra é a detecção de erros de transmissão e as retransmissões; outra ainda é o reagrupamento dos pacotes no destino na ordem em que foram transmitidos – ao que damos o nome de seqüenciação – e muitas outras tarefas, além do gerenciamento de todo o hardware de transmissão.

A realização dessas tarefas é difícil, tem um custo elevado e afasta cada modulo processador (ETD) de seus objetivos primários, que são as aplicações do sistema. De um modo geral, em redes geograficamente distribuídas comutadas por pacotes, isso leva à inclusão de sistemas externos de controle responsáveis pela realização de varias das tarefas mencionadas (e outras). São os ECDs: *Equipamentos de Comunicação de Dados* (ou *Data Communicating Equipments* – DCEs). Equipamentos para concentrar o tráfego interno (denominado nós de comutação ou *Data Switching Equipments* – DSEs) e funcionar como pontos intermediários de restauração dos sinais no interior da rede também são comumente encontrados em redes geograficamente distribuídas.



Em uma **rede geograficamente distribuída comutada por pacotes**, um ECD é, em geral, compartilhado por vários ETDs. O arranjo topológico formado pelos ECDs juntamente com os nós de comutação e as regras de comunicação que executam é o que usualmente chamamos de *sub-rede de comunicação*.

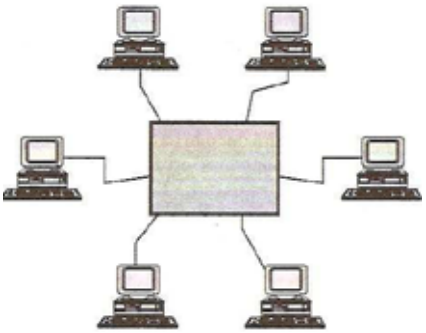
Essas sub-redes são, na sua grande maioria, operadas por empresas especializadas no fornecimento de serviços de comunicação. A topologia final utilizada em redes geograficamente distribuídas pode ser visualizada na imagem acima.

6.6.2 Redes Locais e Metropolitanas

Nas redes LAN e MAN encontramos os seguintes tipos de topologias:

- Topologia em Estrela;
- Topologia em Anel;
- Topologia em Barra;

As características geográficas das redes locais e metropolitanas levam a considerações de custo e tecnologia bastante diferentes das redes de longa distancia. Comentamos na seção anterior que os caminhos alternativos entre os nós da rede eram necessários para aumentar a confiabilidade e desempenho (velocidade efetiva do sistema). Ora, uma forma de aumentarmos a confiabilidade é utilizarmos meios de transmissão com taxas de erro menores; uma forma de melhorarmos desempenho é utilizarmos meios de transmissão de maior velocidade. Em redes locais e metropolitanas, meios de transmissão de alta velocidade, de baixa taxa de erro, de baixo custo e privados podem ser usados. Topologias muitas vezes inviáveis em ambientes geograficamente distribuídos podem ser utilizadas. Examinares a seguir as topologias mais utilizadas nessas redes: estrela, anel e barra.



Uma rede com **topologia em estrela** é ilustrada ao lado. Nesse tipo de topologia cada nó é interligado a um nó central (mestre), através do qual todas as mensagens devem passar. Tal nó age, assim, como centro de controle da rede, interligando os demais nós (escravos). Nada impede que haja comunicações simultâneas, desde que as estações envolvidas sejam diferentes.

Várias redes em estrela operam em configurações onde o nó central tem tanto a função de gerência de comunicação como facilidades de processamento de dados. Em outras redes, o nó central tem como única função o gerenciamento das comunicações. O nó central, cuja função é o chaveamento (ou comutação) entre as estações que desejam se comunicar, é denominado *comutador* ou *switch*.

O arranjo em estrela, evidentemente, é a melhor escolha se o padrão normal de comunicação na rede combinar com essa topologia, isto é, um conjunto de estações secundárias se comunicando com o nó central. Este é, por exemplo, o caso típico das redes de computadores onde o nó central é um sistema de computação que processa informações alimentadas pelos dispositivos periféricos (nós escravos). As situações mais comuns, no entanto, são aquelas em que o nó central está restrito às funções de gerente das comunicações e a operações de diagnóstico.

Redes em estrela podem atual por difusão (broadcasting) ou não. Em redes por difusão, todas as informações são enviadas ao nó central que é o responsável por distribuí-las a todos os nós da rede. Os nós aos quais as informações estavam destinadas compiam-nas e os outros simplesmente as ignoram. Em redes que não operam por difusão, um nó pode apenas se comunicar com outro nó de cada vez, sempre sob controle do nó central.

Redes em estrela não têm necessidade de roteamento, uma vez que concentram todas as mensagens no nó central. O gerenciamento das comunicações por este nó pode ser por chaveamento de pacotes ou chaveamento de circuitos. As redes em estrela podem ainda operar em modo transferência assíncrono (*Assynchronous Transfer Mode* – ATM), como veremos mais adiante no curso. No primeiro caso, pacotes são enviados do nó de origem para o nó central que o retransmite então ao nó de destino no momento apropriado. Já no caso de chaveamento de circuitos, o nó central, baseado em informações recebidas, estabelece uma conexão entre o nó de origem e o nó de destino, conexão esta que exigirá durante toda a conversação. Neste caso, se já existir uma conexão ligando duas estações, nenhuma outra conexão poderá ser estabelecida para esses nós. Redes de chaveamento computadorizadas – CBX (Computerized Branh Exchange) – são exemplos desde último tipo de rede, onde o chaveamento é realizado por um PABX (Private Automatic Branch Exchange).

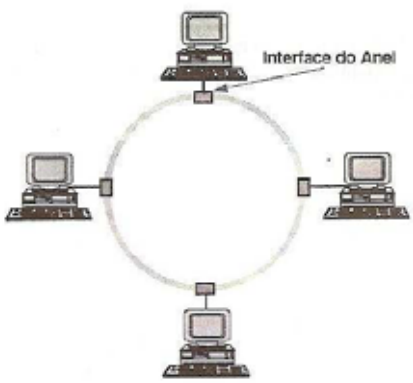
Embora tenhamos incluído a CBX como uma categoria de rede local devido ao fato de tratar de uma alternativa para a interconexão de dispositivos digitais, sua arquitetura e tecnologia são tão diferetnes das demais redes locais de computadores que, frequentemente, não são consideradas como uma rede local de computadores. As CBX são apropriadas tanto para o tráfego de voz quanto para o tráfego de dados entre terminais e entre terminais e computadores. O interesse por esse tipo de redes tem aumentado muito com o desenvolvimento de padrões e pastilhas dedicadas para as chamadas *Redes Digitais com Serviços Integrados* – RDSI (*Integrated Service Digital Network* – ISDN).

Como já mencionamos, o nó central pode realizar funções além das de chaveamento e processamento normal. Por exemplo, o nó central pode realizar a compatibilidade da velocidade de comunicação entre o transmissor e o receptor. Os dispositivos de origem e destino podem até operar com protocolos e/ou conjunto de caracteres diferentes. O nó central atuaria nesse caso como um conversor de protocolos permitindo ao sistema de um fabricante trabalhar satisfatoriamente com um outro sistema de um outro fabricante. Poderia ser também função do nó central fornecer algum grau de proteção de forma a impedir pessoas não autorizadas de utilizar a rede ou ter acesso a determinados sistemas de computação. Outras funções, como operações de diagnóstico de redes, por exemplo, poderiam também fazer parte dos serviços realizados pelo nó mestre.

Confiabilidade é um problema nas redes em estrela. Falhas em um nó escravo apresentam um problema mínimo de confiabilidade, uma vez que o restante da rede ainda continua em funcionamento. Falhas no nó central, por outro lado, podem ocasionar a parada total do sistema. Redundâncias podem ser acrescentadas, porém o custo de tornar o nó central confiável pode mascarar o beneficio obtido com a simplicidade das interfaces exigidas pelas estações secundárias.

Outro problema da rede em estrela é relativo à modularidade. A configuração pode ser expandida até certo limite imposto pelo nó central: em termos de capacidade de chaveamento, número de circuitos concorrentes que podem ser gerenciados e número total de nós que podem ser servidos. Embora não seja freqüentemente encontrado, é possível a utilização de diferentes meios de transmissão para ligação dos nós escravos ao nó central.

O desempenho obtido em uma rede em estrela depende da quantidade de tempo requerido pelo nó central para processar e encaminhar uma mensagem, e da carga de tráfego na conexão, isto é, o desempenho é limitado pela capacidade de processamento do nó central. Um crescimento modular visando o aumento do desempenho torna-se a partir de certo ponto impossível, tendo como única solução a substituição do nó central.



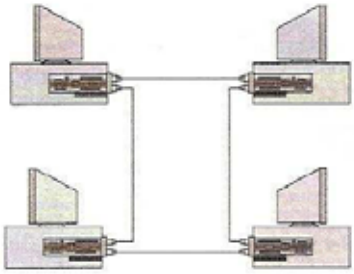
Uma **rede em anel** consiste em estações conectadas através de um caminho fechado. Por motivos de confiabilidade que se tornarão claros ao longo desta seção, o anel não interliga as estações diretamente, mas consiste em uma série de repetidores ligados por um meio físico, sendo cada estação ligada a esses repetidores, conforme apresenta a ilustração ao lado.

Redes em anel são, teoricamente, capazes de transmitir e receber dados em qualquer direção. As configurações mais usuais, no entanto, são unidirecionais, de forma a simplificar o projeto dos repetidores e tornar menos sofisticados os protocolos de comunicação que asseguram a entrega da mensagem ao destino corretamente e em sequência, pois sendo unidirecionais evitam o problema de roteamento. Os repetidores são em geral projetados de forma a transmitir e receber dados simultaneamente, diminuindo assim o retardo de transmissão.

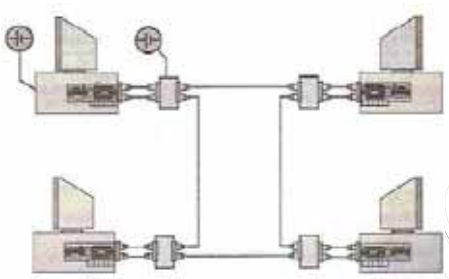
Quando uma mensagem é enviada por um nó, ela entra no anel e circula até ser retirada pelo nó de destino, ou então até voltar ao nó de origem, dependendo do protocolo empregado. No primeiro procedimento, o repetidor deve introduzir um retardo suficiente para o recebimento e armazenamento dos bits de endereçamento de destino da mensagem, quando então poderá decidir se esta deve ou não continuar no anel. No ultimo procedimento, à medida que os bits de uma mensagem vão chegando eles vão sendo despachados, podendo a rede atuar com um retardo de um bit por repetidor. Esse procedimento permite a construção de repetidores mais simples e, por conseqüência, menos susceptíveis a falhas, e de menor custo.

Além da maior simplicidade e do menor retardo introduzido, as redes onde a mensagem é retirada pelo nó de origem permitem mensagens de difusão (broadcast e multicast), isto é, um pacote é enviado simultaneamente para múltiplas estações. Essas redes também possibilitam a determinadas estações receberem mensagens enviadas por qualquer outra estação da rede, independentemente de qual se já o nó de destino. Chamaremos a isto de reconhecimento de *endereçamento promíscuo* ou *modo espião*. Em estações no modo espião podemos, por exemplo, desenvolver programas para observação do trafego dos canais, construir matrizes de trafego, fazer analise de carregamento, realizar isolamento de falhas e protocolos de manutenção, etc.

Topologia em anel requer que cada nó seja capaz de remover seletivamente mensagens da rede ou passá-las à frente para o próximo nó. Isto requer um repetidor ativo em cada nó, e a rede poderá ser mais confiável do que esses repetidores. Uma quebra em qualquer dos enlaces entre os repetidores vai para toda a rede até que o problema seja isolado e um novo cabo instalado. Falhas no repetidor ativo também podem causar a parada total do sistema.

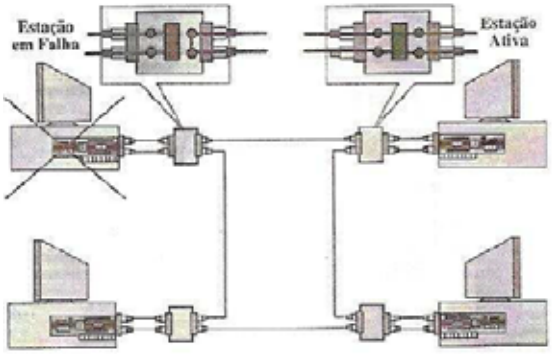


Se os repetidores fizessem parte do hardware específico e interno de cada estação conectada à rede, como na imagem acima, a vulnerabilidade seria ainda maior: os repetidores estariam susceptíveis a falhas no equipamento ou à própria falta de alimentação elétrica da estação. Por esse motivo, os repetidores são alimentados e mantidos separados do hardware da estação como ilustrado abaixo.

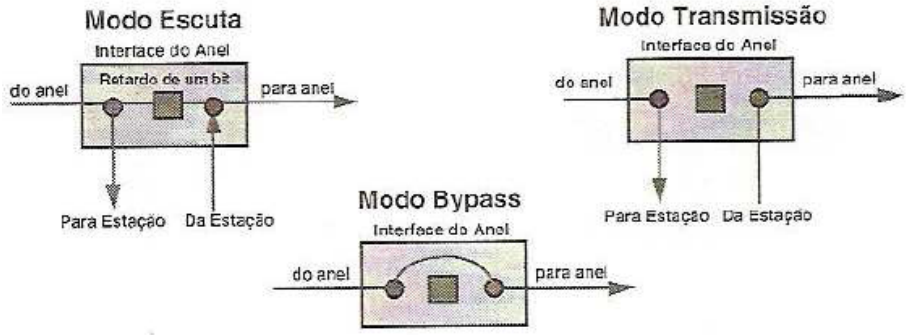


Uma solução parcial para o problema de falha no repetidor consta em prover cada um deles de um relé que pode removê-lo mecanicamente da rede em caso de falha, como apresentado na imagem da direita.

Essa remoção pode ser impossível se os repetidores imediatamente posteriores e anteriores ao repetidor com falha estiverem a uma distância maior do que o limite exigido pelo meio de transmissão para a interconexão de dois nós – devido ao problema da atenuação, que veremos também mais adiante.



Provida de relé, a interface possui três modos ou estados de funcionamento: escuta, transmissão ou bypass.



No estado de Escuta, cada bit que chega ao repetidor é retransmitido com o menor retardo possível (o ideal é da ordem de um bit), apenas suficiente para a realização das seguintes funções:

1. Análise do fluxo de dados para procura de determinados padrões de bits, como por exemplo: endereços das estações conectadas ao repetidor, permissão de controle, etc.
2. Em caso de pacotes endereçados a estação e detectados em (1), deve ser realizada a cópia de cada bit do fluxo de entrada e feito o envio à estação, ao mesmo tempo que esses bits são retransmitidos.
3. Modificação de bits do fluxo de entrada para a retransmissão, necessária em certas estratégias de controle de erros.

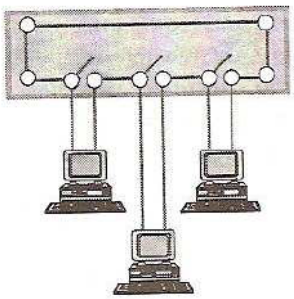
Quando uma estação adquire o direito de acesso à rede através de algum esquema de controle e têm dados a transmitir, a interface entre no estado de transmissão. Nesse estado os bits recebidos da estação são transmitidos pela interface, que durante este período pode receber um fluxo de bits do anel, cujo significado vai exigir dois tratamentos distintos, para duas situações diferentes:

1. Os bits que a interface recebe do anel podem ser da própria estação que está transmitindo. Isto ocorrerá se o retardo do anel for menor que o tempo de transmissão de uma mensagem, quando os bits iniciais transmitidos já estarão retornando à própria estação antes do final da transmissão. Nesse caso, a interface retorna os bits recebidos à estação de modo que ela possa checá-los como uma forma de reconhecimento, ou simplesmente os descarta.
2. Alguns esquemas de controle permitem que mais de uma mensagem circule no canal ao mesmo tempo. Se a interface, enquanto estiver transmitindo, receber bits que não foram os originados por ela própria (caso 1), os bits recebidos devem ser armazenados para posterior transmissão (pois, nesse caso, trata-se de uma mensagem gerada por outra estação, que deverá ser retransmitida). Em caso contrário, age como em 1.

Os dois estados, transmissão e recepção, são suficientes para a operação do anel. O terceiro estado – o de *bypass* – é utilizado para aumentar a confiabilidade da rede, conforme já discutimos. Nesse estado, um relé é ativado de forma que o fluxo de dados de entrada passe pela interface diretamente para a saída, sem nenhum retardo ou regeneração. Esse estado traz dois benefícios à rede. O primeiro é a solução parcial do problema de confiabilidade já discutido. O segundo é a melhora no desempenho através da eliminação do retardo introduzido na rede por estações que não estão ativas.

Outras melhoras na topologia em anel foram propostas e realizadas, como a introdução de caminhos alternativos, duplos anéis e etc. Experiências práticas sugerem que a topologia pode ser feita suficientemente confiável de forma que a possibilidade de falhas possa ser praticamente ignorada. É claro que o custo pode tornar a rede confiável proibitiva para certas aplicações. Analisemos resumidamente algumas dessas melhoras.

A primeira delas é a introdução de concentradores (*ring wiring concentrators*), também denominados *hubs*, como ilustrado na figura ao lado.



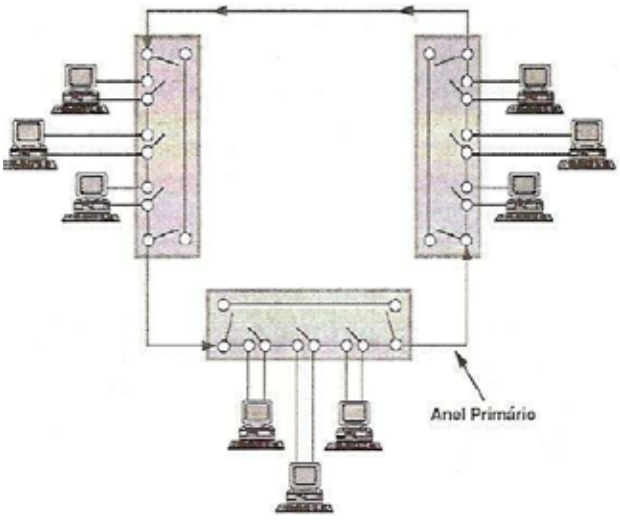
Inicialmente esses concentradores eram apenas elementos passivos que permitiam a concentração de todo o cabeamento utilizado e possuíam um mecanismo de relés que, acionado externamente, permitia o isolamento de estações em falha. Mais tarde eles passaram a ser utilizados como concentradores dos repetidores do anel (concentradores ativos).

Tal técnica tem várias vantagens. O isolamento de falhas se torna mais simples porque existe um ponto de acesso central para o sinal. Sem o concentrador, quando um repetidor ou um enlace falha, a localização da falha requer uma busca através de todo o anel, exigindo o acesso a todos os locais que contêm repetidores e cabos.

Outra vantagem do concentrador é a possibilidade de adição de novas estações sem a parada total da rede, uma vez que novos repetidores podem ser ativados no concentrador, sem para a rede, por meio da utilização de relés.

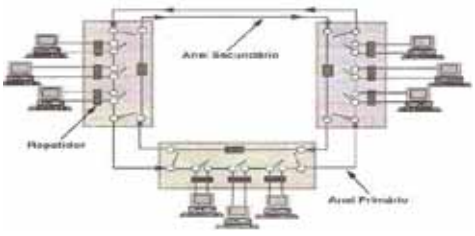
A modularidade de uma rede em anel é bastante elevada devido ao fato de os repetidores ativos regenerarem as mensagens. Redes em anel podem atingir grandes distâncias (teoricamente o infinito). Existe no entanto uma limitação prática do número de estações em um anel. Esse limite é devido aos problemas de manutenção e confiabilidade, citados anteriormente, e ao retardo cumulativo do grande número de repetidores.

A imagem ao lado apresenta um anel formado pela interconexão de concentradores. Devemos sempre lembrar que a distancia entre dois concentradores não deverá ultrapassar o limite máximo permitido sem regeneração do sinal.

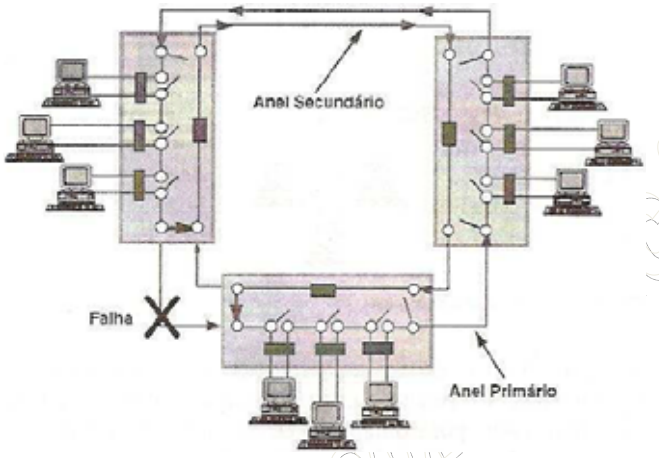


Embora a utilização de relés permita a rápida recuperação de algumas falhas nos repetidores, existem outras falhas que podem temporariamente parar toda a rede, como por exemplo falhas nos segmentos entre os concentradores.

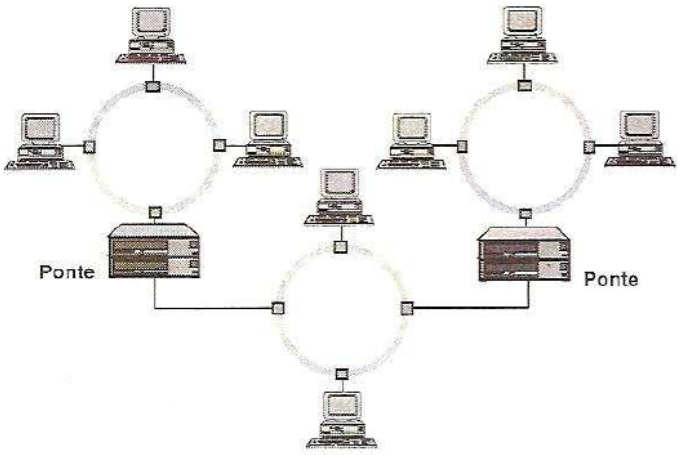
Uma solução para o problema seria a utilização de caminhos alternativos: duplo anel, triplo anel:



No duplo anel, um dos anéis é o anel principal e o outro é acionado somente em caso de falhas, sendo denominado anel secundário ou anel de backup. O anel de backup tem sua orientação definida no sentido contrario ao do anel principal. A figura abaixo mostra como o anel de backup entra em funcionamento no caso de uma falha em um segmento de cabo.



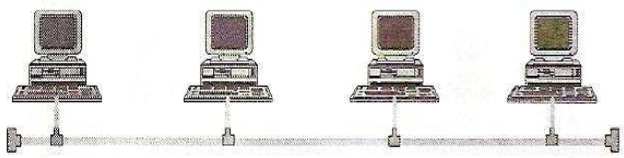
Outra solução para aumentar a confiabilidade de uma rede em anel seria considerar a rede local como consistindo em vários anéis, e o conjunto dos anéis conectados por *pontes* (*bridges*). A ponte encaminha os pacotes de dados de uma sub-rede a outra com base nas informações de endereçamento. Do ponto de vista físico, cada anel operaria independentemente. Ressaltamos então dois fatos. Primeiro, uma falha em um anel vai para somente aquela porção da rede.



Uma falha na ponte não impede o trafego intra-rede. Segundo, múltiplos anéis podem ser empregados para a obtenção de um maior nível de desempenho.

Como vimos, os maiores problemas com topologias em anel são sua vulnerabilidade a erros e pouca tolerância a falhas. Qualquer que seja o controle de acesso empregado, ele pode ser perdido por falhar e pode ser difícil determinar com certeza se esse controle foi perdido ou decidir qual nó deve recriá-lo. Erros de transmissão e processamento podem fazer com que uma mensagem continue eternamente a circular no anel. Embora não seja essencial do ponto de vista de projeto, uma estação monitora tem-se revelado essencial, na prática, na maioria dos anéis. A função primordial desta estação monitora é a de contornar os problemas mencionados. Outra de suas funções é iniciar o anel, enviar mensagens de teste e diagnósticos e outras tarefas de manutenção. A estação monitora pode ser uma estação de dedicada ou então uma estação qualquer da rede que assuma em determinado tempo tais funções, como veremos com mais detalhes posteriormente.

Por serem geralmente unidirecionais, redes com topologias em anel são ideais para utilização de fibra ótica. Existem algumas redes que combinam seções de diferentes meios de transmissão sem nenhum problema.



A última, a **topologia em barra**, é aquela onde todas as estações (nós) se ligam ao mesmo meio de transmissão. Ao contrário das outras topologias que discutimos até aqui, que são configurações ponto a ponto (isto é, cada enlace físico de transmissão conecta apenas dois dispositivos), a topologia em barra tem uma configuração multiponto.

Nas redes em barra comum cada nó conectado à barra pode ouvir todas as informações transmitidas, similar às transmissões de radiodifusão. Esta característica vai facilitar as aplicações com mensagens do tipo difusão (mensagens globais) além de possibilitar que algumas estações possam trabalhar no que chamamos de endereçamento promíscuo ou modo espião.

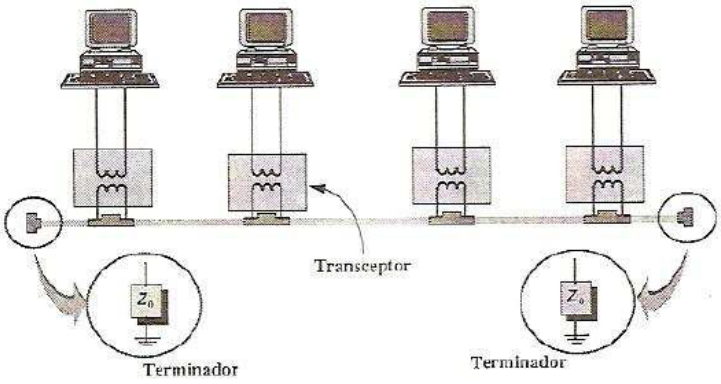
Existe uma variedade de mecanismos para o controle de acesso à barra, que pode ser centralizado ou descentralizado. A técnica adotada para cada acesso à rede (ou à banda de freqüência de rede no caso de redes em banda larga, como veremos adiante) é uma forma de multiplexação no tempo. Em um controle centralizado, o direito de acesso é determinado por uma estação especial da rede. Em um ambiente de controle descentralizado, a responsabilidade de acesso é distribuída entre todos os nós.

Ao contrário da topologia em anel, as topologias em barra podem empregar interfaces passivas, nas quais as falhas não causam a parada total do sistema. Relógios de prevenção (*watch-dog timers*) em cada transmissor devem detectar e desconectar o nó que falha no modo de transmissão (nó que não pára de transmitir). A confiabilidade desse tipo de topologia vai depender em muita da estratégia de controle. O controle centralizado oferece os mesmos problemas de confiabilidade de uma rede em estrela, com o atenuante de que, aqui, a redundância de um nó pode ser outro nó comum da rede. Mecanismos de controle descentralizados semelhantes aos empregados na topologia em anel podem também ser empregados neste tipo de topologia, acarretando os mesmos problemas quanto à detecção da perde do controle e sua recriação.

A ligação ao meio de transmissão é um ponto crítico no projeto de uma rede local em barra comum. A ligação deve ser feita de forma a alterar o mínimo possível as características elétricas do meio. O meio, por sua vez, deve terminar em seus dois extremos por uma carga igual a sua impedância característica, de forma a evitar reflexões espúrias que interfiram no sinal transmitido.

A ligação das estações ao meio de comunicação é realizada através de um transceptor (transmissor/receptor), que tem como funções básicas transmitir e receber sinais, bem como reconhecer a presença destes sinais no meio. O transceptor se liga à barra através de um conector, que é responsável pelo contato elétrico com os condutores da barra.

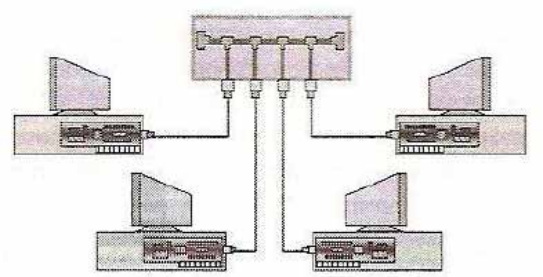
Ligações ao meio de transmissão geram descontinuidade de impedância, causando reflexões.



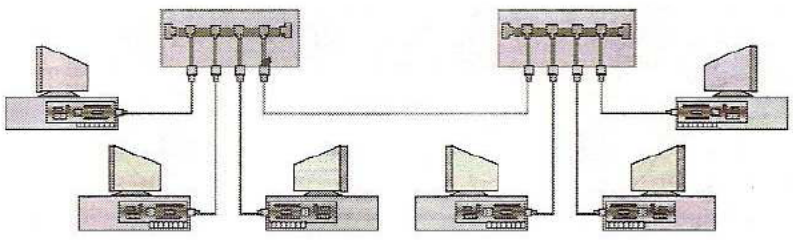
Assim, o transceptor deve apresentar uma alta impedância para o cabo, de forma que sua ligação a este altere o mínimo possível as características de transmissão. Devido a isto, o transceptor deve ser localizado perto do cabo (uma distância grande do cabo impediria a obtenção de uma alta impedância), a uma distância de alguns poucos centímetros, como mostra a figura ao lado.

O poder de crescimento, tanto no que diz respeito à distância máxima entre dois nós da rede quanto ao número de nós que a rede pode suportar, vai depender do meio de transmissão utilizado, da taxa de transmissão e da quantidade de ligações ao meio. Conforme se queira chegar a distâncias maiores que a máxima permitida em um segmento de cabo, repetidores serão necessários para assegurar a qualidade do sinal. Tais repetidores, por serem ativos, apresentam um ponto de possível diminuição da confiabilidade da rede.

Assim como em redes em anel, a utilização de concentradores (hubs) irá facilitar a localização e o isolamento de falhas, bem como permitir a inserção de novas estações na barra sem a parada do sistema (caso existam entradas livres no hub). A figura ao lado mostra o hub de uma rede em barra.



Hubs podem ser interconectados como forma de expansão do tamanho da rede, conforme ilustrado na imagem abaixo:



O desempenho de um sistema em barra comum é determinado pelo meio de transmissão, número de nós conectados, controle de acesso, tipo de tráfego e outros fatores. Por empregar interfaces passivas (sem repetidores), que não exigem armazenamento local de mensagens, topologias em barra não vão degradar o retardo de transferência, que, contudo, pode ser altamente dependente do protocolo de acesso utilizado.

6.6.3 Comparações entre as Topologias

De forma geral podem realizar as seguintes comparações entre as topologias vistas:

Topologia / Características	ESTRELA	ANEL	BARRA COMUM	GRAFOS
Simplicidade Funcional	a melhor de todas	razoável	razoável, um pouco melhor do que o anel	extremamente complexa
Roteamento	inexistente	inexistente no anel unidireciona, simples nos outros tipos	inexistente	bastante complexo
Custo de Conexão	alto (incluindo o custo do nó central)	baixo para médio	baixo	muito alto
Crescimento Incremental	limitado a capacidade do nó central	teoricamente infinito	alto	alto
Aplicação Adequada	aquelas envolvendo processamento central de todas as mensagens	sem limitação	sem limitação	sem limitações
Desempenho	baixo, todas as mensagens têm de passar pelo nó central	auto, possibilidade de mais de uma mensagem ser transmitida ao mesmo tempo	médio	alto. pode se adaptar ao volume de tráfego existente
Confiabilidade	pouca confiabilidade	boa, desde que sejam tomados cuidados adicionais	a melhor de todas. interface passiva com o meio	boa, devido a existência de caminhos alternativos
Retardo de Transmissão	médio	baixo, podendo chegar a não mais que 1 bit por nó	o mais baixo de todas	alto
Limitação Quanto ao Meio de Transmissão	nenhuma. ligação ponto a ponto	nenhuma. ligação ponto a ponto	por ter a ligação multiponto sua ligação ao meio de transmissão pode ser de custo elevado, como é o caso da fibra ótica	nenhuma. ligação ponto a ponto

Tipos de Topologias	Ponto Positivos	Pontos Negativos
Topologia Estrela	• É mais tolerante a falhas • Fácil de instalar usuários • Monitoramento centralizado	• Custo de Instalação maior porque recebe mais cabos
Topologia Anel (Token Ring)	• Razoavelmente fácil de instalar • Requer menos cabos • Desempenho uniforme	• Se uma estação para todas param • Os problemas são difíceis de isolar.
Topologia Barramento	• Simples e fácil de instalar • Requer menos cabos • Fácil de entender	• A rede fica mais lenta em períodos de uso intenso. • Os problemas são difíceis de isolar.

6.7 HUBS E SWITCHES

A topologia de uma rede irá determinar, em parte, o método de acesso utilizado. Métodos de acesso são necessários para regular o acesso a meios físicos compartilhados. Assim, costuma-se associar os métodos de acesso (que estudaremos na próxima competência) às topologias utilizadas. Como vimos ao longo deste capítulo, a instalação física das redes tem sofrido uma forte tendência na direção da utilização de hubs, o que, fisicamente, corresponde à implantação de uma topologia em estrela. Essa tendência é explicada, basicamente, pela crescente necessidade de melhorar o gerenciamento e a manutenção nessas instalações. O maior problema da topologia em estrela, como mencionado, é a sua baixa confiabilidade dada a presença de um elemento central no qual as falhas provocam a parada total do sistema. Porém, os avanços da eletrônica já permitem, hoje, que se construam equipamentos de alta confiabilidade, viabilizando esse tipo de topologia.

A utilização de hubs, no entanto, não exige, necessariamente, que as interfaces das estações com a rede a percebam como uma topologia em estrela. Do ponto de vista da interface das estações com a rede, o funcionamento se dá como em uma barra ou em um anel, com os seus respectivos métodos de acesso. Note porém, que a implementação física, interna nos hubs, pode ser qualquer uma desde que essa interface seja preservada.

Pelo que acabamos de apresentar, podemos diferenciar dois tipos de topologias: uma topologia lógica, que é aquela observada sob o ponto de vista das interfaces das estações com a rede (que inclui o método de acesso), e uma topologia física, que diz respeito ao layout físico utilizado na instalação da rede.

A demanda por maiores taxas de transmissão e melhor utilização dos meios físicos, aliados à evolução da microeletrônica, começou a alterar a construção desses equipamentos concentradores. A partir do momento em que as estações estão ligadas a um elemento central, no qual a implementação interna é desconhecida mas a interface é coerente com as estações, é possível pensar que esses elementos podem implementar arquiteturas que não utilizam apenas um meio compartilhado, mas sim possibilitam a troca de mensagens entre várias estações simultaneamente. Dessa forma, estações podem obter para si taxas efetivas de transmissão bem maiores do que as observadas anteriormente. Esse tipo de elemento central é denominado (assim como na topologia em estrela) ***switch***.

Seguir essa tendência utilizando-se dos métodos de acesso para meios compartilhados impõe limitações muito grandes às taxas de transmissão que se pode atingir, muito embora tenha sido uma necessidade de mercado manter as interfaces anteriormente padronizadas. Mas a evolução natural, como não poderia deixar de ser, veio com a criação de novas interfaces de acesso que permitiram que taxas de transmissão bem maiores fossem utilizadas. Redes ATM, como veremos adiante, baseiam-se na presença de switches de grande capacidade de comutação que permitem taxas de transmissão que podem chegar à ordem de Gigabits/s.

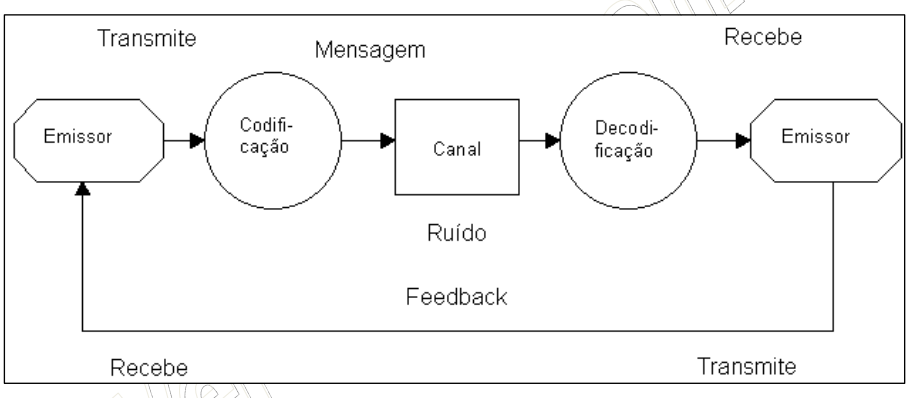
Assim, a topologia em estrela, tanto física quanto logicamente, retoma seu lugar no mundo das redes de computadores. Veremos ao longo deste curso os diversos métodos de acesso utilizados em redes com topologia (lógica) em barra e em anel, e os padrões que foram definidos para essas redes, incluindo as opções de topologia física.

7 COMPETÊNCIA 2 – FUNDAMENTOS DE COMUNICAÇÃO DIGITAL

7.1 TRANSMISSÃO DE INFORMAÇÃO

A transmissão de informação através de sistemas de comunicação pressupõe a passagem de sinais através dos meios de comunicação que compõe as redes. As propriedades físicas dos meios de transmissão e as características dos sinais transmitidos apresentam uma série de questões tecnológicas que influenciam na construção e no projeto de redes de computadores. Neste capítulo, apresentaremos os principais conceitos envolvidos na codificação e transmissão da informação.

O processo de comunicação compreende a transmissão de informação e de significados. Se não há transmissão de informação ou de significado, não há comunicação. Ele envolve a seleção dos assuntos de comunicação, a codificação desta informação, a transmissão da informação codificada e o movimento desta transmissão através dos canais de comunicação para o receptor, que então detecta a informação transmitida; isto é, decodifica a transmissão e seleciona os assuntos comunicativos que são mais importantes para ele.



Em qualquer processo de comunicação, sempre há os seguintes elementos: emissor, receptor, mensagem, canal de comunicação, ruídos e feedback. O emissor trasmite uma mensagem, por algum meio, para um destinatário ou receptor. Antes de transmitir, a fonte codifica a mensagem, convertendo-a em símbolos: idioma, sons, letras, números e outros tipos de sinais. A mensagem segue por um canal, ou meio de comunicação: conversação, telefonema, e-mail, memorando ou outro. Na outra ponta da linha, o receptor decodifica a mensagem, desde que esteja usando o mesmo sistema de símbolos do emissor. A mensagem é, então, interpretada pelo receptor.

A comunicação é o ato de transmitir informação. Ao transmitir informação esperamos preservar o seu significado, recuperar o seu entendimento para permitir a sua manipulação. Um processo de comunicação admite a existência de um código ou linguagem capaz de representar informações através de símbolos compreensíveis para as parte envolvidas. A linguagem verbal é certamente a mais conhecida e utilizada pelo homem.

Quando conversamos, participamos num processo contínuo de conversação das nossas idéias em mensagens numa linguagem verbal, que pode ser transmitida através de sinais acústicos com ajuda das cordas vocais. Os sistemas de comunicação, aqui tratados, utilizam em geral ***sinais*** ou ondas eletromagnéticas que seguem através de meios físicos de comunicação.

Sinais nada mais são do que ondas que se propagam através de alguns meios físicos, seja ele através de ar, um par de fios, etc. Os sinais podem possuir, por exemplo, amplitude que varia ao longo do

tempo correspondendo à codificação da informação transmitida. Os sinais podem, assim, ser representados como uma função do tempo.

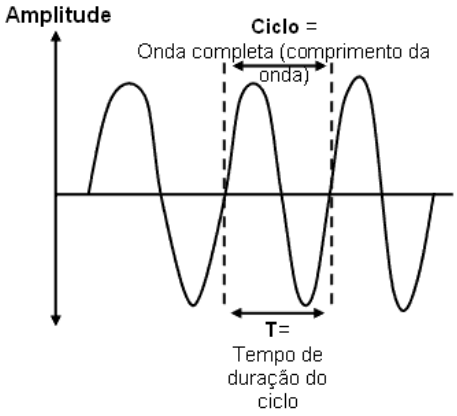
O contexto no qual empregamos os termos sinal e informação é que os diferenciam. *Informações* estão, em geral, associadas às idéias ou aos dados manipulados pelos agentes que as criam, manipula e processam. *Sinais*, por outro lado, correspondem à materialização específica dessas informações utilizadas no momento da transmissão.

Os termos *analógico* e *digital* correspondem, de certa forma, à variação *contínua* e *discreta* respectivamente. Estes termos são freqüentemente usados no contexto das comunicações de dados para qualificar tanto a natureza das informações quanto a característica dos sinais utilizados para a transmissão através dos meios físicos.

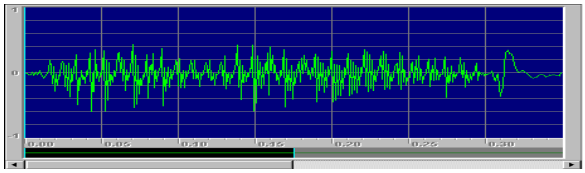
Computadores, por exemplo, são equipamentos que armazenam, processam e codificam informações em bits que correspondem a dois níveis *discretos* de tensão ou corrente, representando os valores lógicos "0" e "1". Chama-se esse tipo de informação de ***digital***. Já informações geradas por fontes sonoras apresentam variações *contínuas* de amplitude, a que denominamos de ***analógica***.

7.2 TIPOS DE TRANSMISSÃO

De forma análoga a que se procedeu a respeito da natureza da informação, podemos classificar em dois os tipos de sinais gerados para a transmissão: *sinais analógicos* e *sinais digitais*.



Sinais analógicos variam continuamente no tempo, como podemos observar na figura ao lado (exemplo de 01 onda) e na figura abaixo (exemplo de um conjunto de ondas). São todos aqueles que ao longo e um intervalo de tempo pré-estabelecido, poderão atingir qualquer valor dentro de um determinado limite. São impulsos sob forma de *Ondas Senoidais.*



A *Onda Senoidal* possui um padrão que se repete:

- Padrão que se repete chamado: ***ciclo;***
- Cada ciclo demora um determinado tempo para ocorrer, chamado de: período ***T;***
- O nº de vezes que o ciclo se repete por segundo: *frequência*, medida em Hertz (Hz=ciclos por segundo);
- A amplitude da onda é a sua altura, medida em Volts no caso de ondas elétricas. O comprimento da onda é dado em metros, dividindo a velocidade da luz pela frequência da onda. Podemos sabê-lo através:

x metros = Velocidade da luz m/seg / frequência da onda

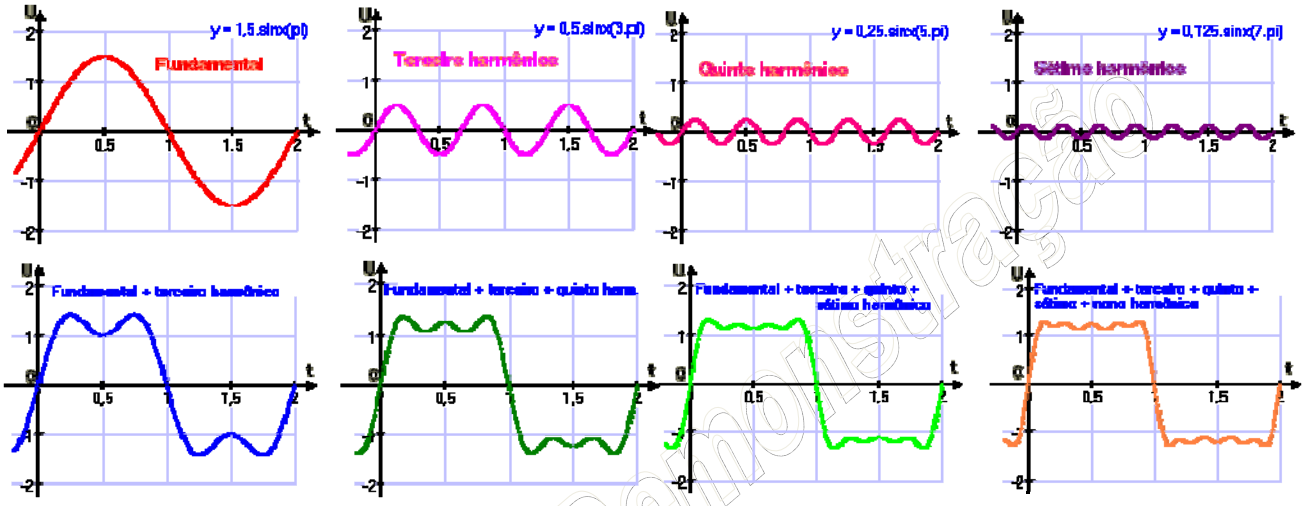
Exemplo:

Qual comprimento de onda em 20 Mhz de freqüência? x= 300.000.000m/seg / 20.000.000Hz

x= 15 metros

Para sabermos qual o **comprimento de onda** numa determinada frequência, basta dividirmos a velocidade de propagação da onda eletromagnética no vácuo (300.000.000 m/s) pela frequência (em Hertz).

As variações constantes da onda senoidal (comprimento, freqüência e período) produzem sinais analógicos de diferentes formatos, constituídos por diversas ondas senoidais, ou mais conhecidas como harmônicas:



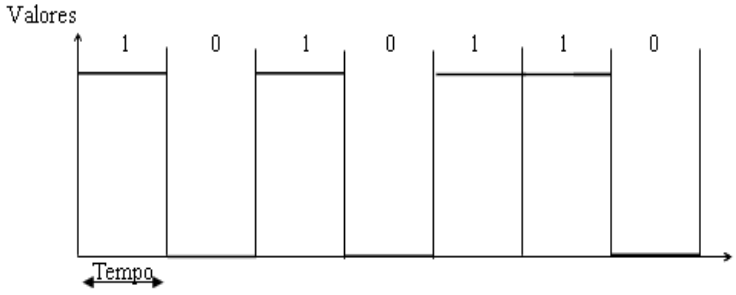
Estas são as ondas básicas para as operações envolvendo funções periódicas.

Sendo o sinal analógico uma onda que varia continuamente e é transmitida por diversos meios, ela está mais sujeita a distorções, atenuações e ruídos ao longo da sua transmissão.

Os sistemas telefônicos analógicos, quando usados para a comunicação de dados, são muito limitados, principalmente no que diz respeito à *largura de banda* (velocidade de comunicação), além de estarem sujeitos à distorção do sinal se a comunicação for realizada através de longa distâncias.

O baixo custo é uma das principais vantagens do uso de sistemas de transmissão telefônicos, no entanto não é um meio adequado para a transmissão de dados, principalmente devido à baixa velocidade. Além disso, a qualidade da transmissão tendo a piorar quando maior for a distância entre os nós.

O sinal digital caracteriza-se pela presença de ***pulsos*** nos quais a amplitude é fixa, como apresentado na figura ao lado. O sinal é construído através de uma sequência de intervalos de tamanho fixo iguais a ***T*** segundos, chamado intervalos de sinalização, durante os quais a amplitude do sinal permanece fixa, caracterizando um dos símbolos digitais transmitidos.



É importante que se entenda que qualquer tipo de informação (seja analógica ou digital) pode ser transmitida através de um sinal analógico ou digital. Um sinal de voz analógico, por exemplo, pode ser amostrado, quantizado e o resultado dessa quantização, codificado em um sinal digital para transmissão. A transmissão de informações digitais através de sinais analógicos também é possível; técnicas de modulação transformam sinais digitais em sinais que apresentam variação continua de amplitude.

Atualmente, a maior parte das tecnologias de rede locais (LAN) os meios de transmissão mantêm os dados em formato digital, enquanto que nas redes de longa distância (WAN), com o uso das linhas telefônicas, os sinais são transmitidos em formato analógico (a exceção das RDSI), e trabalha diretamente com transmissão digital utilizando modems digitais, efetuando técnicas de modulação.

A modulação é o processo de converter dados no formato digital (ondas quadradas) para o analógico (ondas senoidais), o processo inverso é denominado demodulação. O equipamento responsável por esses processos é conhecido por Modem (Modulador-Demodulador).

Os modems digitais são necessários porque o sinal digital possui um alcance pequeno. A solução é passar por uma modulação com uma *portadora (Carrier)* mais adequada ao meio de transmissão. A portadora é uma onda utilizada para transportar dados entre computadores.

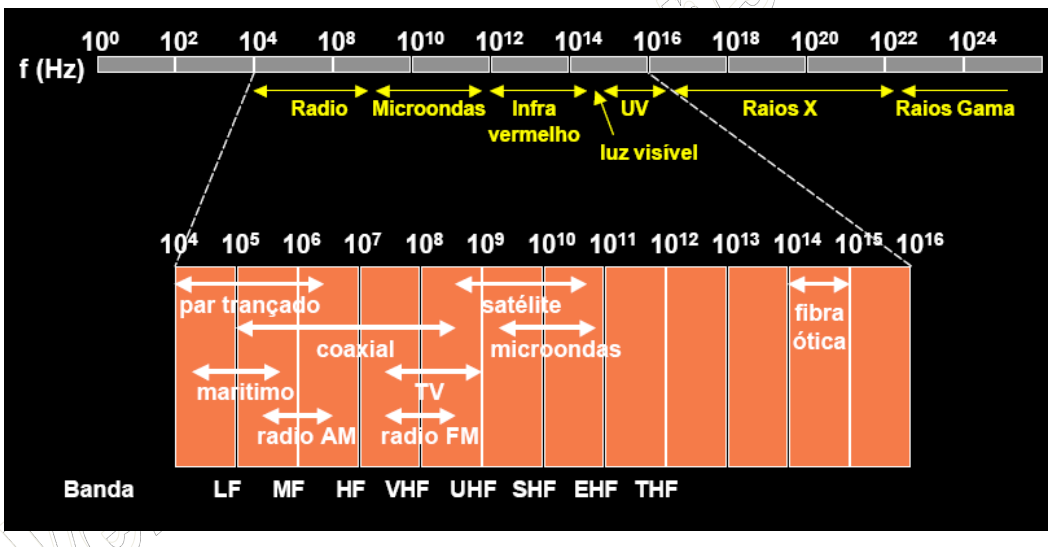
7.3 BANDA PASSANTE E LARGURA DE BANDA

Chama-se **banda passante** o conjunto contínuo de valores de freqüência que podem ser assumidos por um sinal elétrico sem que este seja atenuado ao passar por um filtro. Informalmente, diz-se são as freqüências que "passam" pelo filtro.

O valor de freqüência, medido em Hertz, a partir do qual o sinal não "passa" pelo filtro é chamado de freqüência de corte. Idealmente, sinais com freqüência além ou aquém da(s) freqüência(s) de corte do filtro seriam atenuados a zero. Na prática, entretanto, adota-se o critério de **meia potência**: é (são) considerada(s) freqüência(s) de corte aquelas em que a potência do sinal é atenuada à metade da original.

Dois filtros dados podem ter a mesma largura de banda, digamos 3kHz, mas bandas passantes diferentes; por exemplo, um com banda passante de 1kHz a 4kHz, o outro de 40kHz a 43kHz.

Compreendemos melhor o termo banda passante analisando o espectro de frequência do nosso dia a dia:



No gráfico temos uma representação do espectro de frequência que vai de 0Hz até aproximadamente 1THz. Dentro de cada intervalo de frequencia (banda passante) existe uma aplicação técnica. Quanto mais as tecnologias forem evoluindo maiores serão os usos em frequências mais altas. Quanto mais alta a frequência maior a necessidade de velocidade de processamento dos circuitos intergrados (chips DSP – Digital Signal Processor, desenvolvidos na década de 70 exclusivamente para o processamento de sinais digitais), que atualmente operam apenas na faixa dos Ghz.

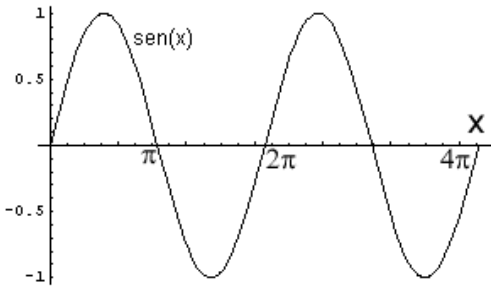
Cada intervalo possui tamanhos diferentes, o que chamamos de largura de banda, ou seja, na banda passante do Radio temos uma largura de banda de aproximadamente $10^8 – 10^4 = 10^4$.

Um dos primeiros grandes estudiosos da matemática que analisou em detalhes as aplicações do espectro de frequência foi o cientista francês do século XIX Jean Fourier, que além de várias teorias e demonstrações, provou que qualquer sinal período, expresso como uma função do tempo g(t), com período T_0, pode ser considerado como uma soma (possivelmente infinita) de senos e cossenos de diversas frequencias. A essa soma, dá-se o nome de Série de Fourier, que pode ser presentada como:

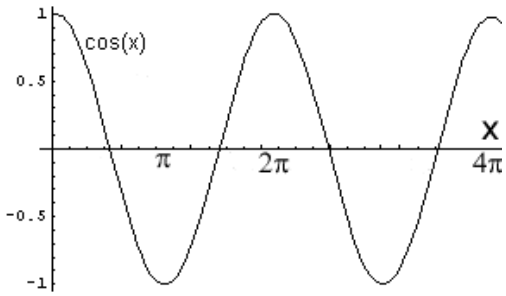
$$g(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \operatorname{sen}(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$

Fourier começou sua carreira publicando diversas aplicações com as séries trigonométricas, desenvolveu sua teoria ao estudar a propagação do calor em corpos sólidos, e afirmava que a forma mais simples de uma onda (calor) é uma função senoidal.

Para explicar a Série de Fourier, e sua importância, inicialmente precisamos recordar um pouco sobre funções períodicas, senos e cossenos:



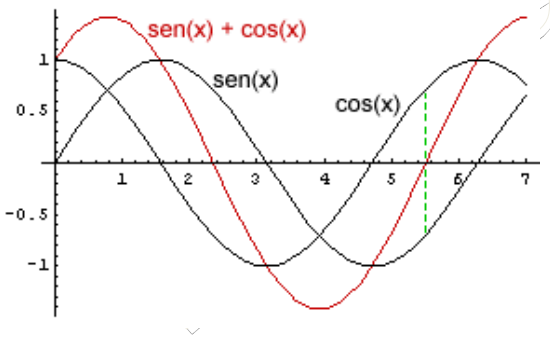
A função seno se repete a cada período de 2π . O valor máximo da função, chamado de AMPLITUDE, é 1.



A função cosseno também é periódica, com o mesmo período e amplitude que o seno, mas é deslocada de π /2 em relação ao seno.

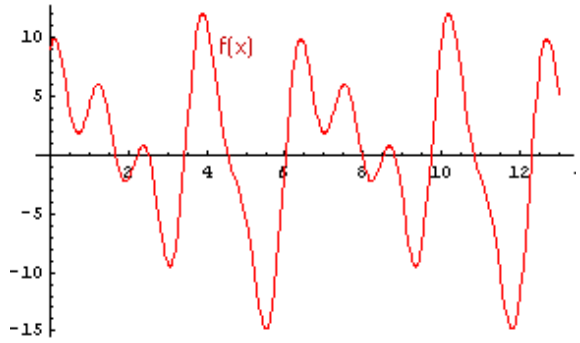
As funções seno e cosseno diferem na FASE e a diferença de fase entre elas é de π/2.

A idéia de Fourier começa com o princípio básico de que essas funções podem ser somadas:



Na figura vemos a soma (curva em vermelho) das funções sen(x) e cos(x). Essa curva é obtida traçando-se, em cada ponto x, a soma dos valores de sen(x) e cos(x) nesse ponto. Por exemplo, o ponto da curva na região x=5,5 é zero pois o valor de sen(x) é igual e de sinal oposto ao valor de cos(x) nesse ponto. Verifique a situação para outros pontos da curva para treinar pois as séries de Fourier são composições de muitas curvas tipo seno e cosseno, como veremos.

Porém, no dia a dia, as ondas/sinais/funções são bem mais complicadas que uma senóide. Veja o exemplo da função f(x) mostrada abaixo. Essa curva também é periódica, mas, não é apenas um seno ou um cosseno. Como achar uma função matemática que descreva uma curva como essa?

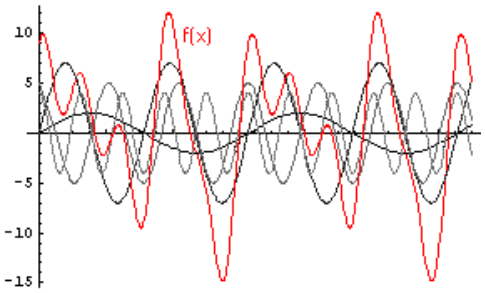


Fourier descobriu que: ***Qualquer função periódica, por mais complicada que seja, pode ser representada como a soma de várias funções seno e cosseno com amplitudes, fases e períodos escolhidos convenientemente.***

Em resumo, qualquer função f(x) pode, segundo Fourier, ser escrita na forma da soma de uma série de funções seno e cosseno da seguinte forma geral:

$$f(x) = a_0 + a_1\,\mathrm{sen}(x) + a_2\,\mathrm{sen}(2x) + a_3\,\mathrm{sen}(3x) + \ldots + b_1\cos(x) + b_2\cos(2x) + b_3\cos(3x) + \ldots$$

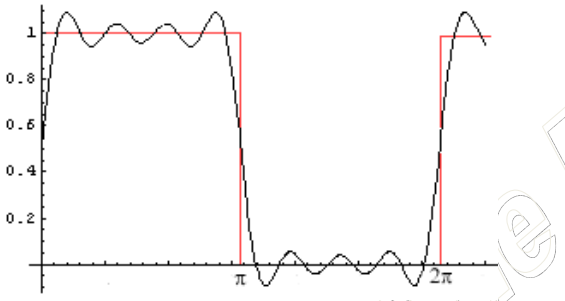
Exemplo:



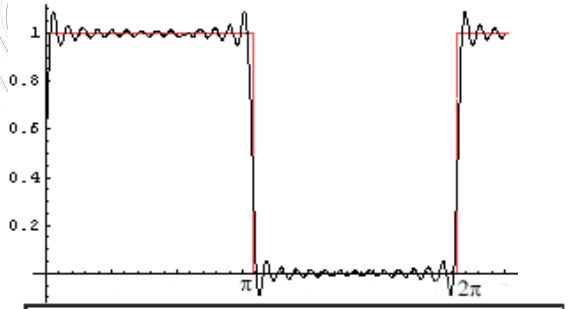
Matematicamente, a decomposição da função f(x) na curva acima é a seguinte:

$$f(x) = 2\,\mathrm{sen}(x) + 7\,\mathrm{sen}(2x) + 5\cos(3x) + 4\cos(5x)$$

A teoria das Séries de Fourier apresenta diversas aplicações práticas. Para as redes de computadores, e mais especificamente para a área de comunicação de dados, representa a forma como transformar uma onda analógica em digital, e vice-versa, vejamos como funciona:



A figura acima mostra um gráfico da onda quadrada juntamente com o gráfico da expansão com os primeiros 5 termos (harmônicas) da série de Fourier



A figura mostra a onda quadrada e sua expansão com os 15 primeiros termos da série de Fourier. Quanto maior o número de termos na expansão, melhor a aproximação com a forma da função original.

Série de Fourier para a onda quadrada é:
f(x) = 1/2 + (2/) sen(x) + (2/(3)) sen(3x) + (2/(5)) sen(5x) + (2/(7)) sen(7x) + ...

Ou seja, quanto mais termos harmônicos procurarmos através da Série de Fourier, melhor a aproximação com a forma da onda digital. Fourier continuou os seus estudos e desenvolveu novas teorias sobre suas Séries, um desses estudos é a Transformada de Fourier, que permite determinar a banda passante de um sinal e com isso definirmos as aplicações segundo suas freqüências e desenvolver processadores capazes de se comunicar dentro da banda passante da aplicação. Uma das conclusões possíveis de se extrair sobre a Transformada de Fourier, é que a banda passante é a largura de banda mínima capaz de garantir que o receptor ainda recupere a informação digital originalmente transmitida.

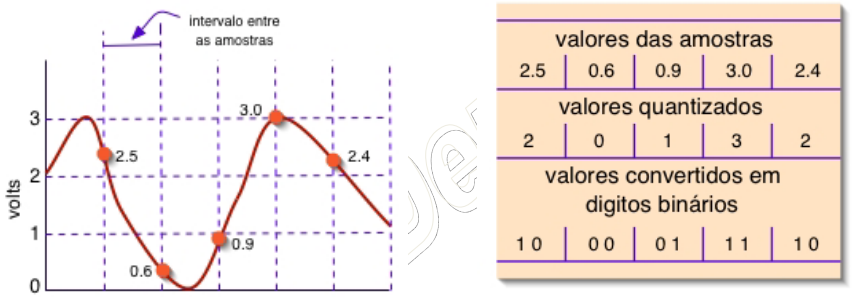
A largura de banda do sinal digital, por sua vez, depende do tamanho T dos pulsos (o intervalo de sinalização); em outras palavras: depende da velocidade em bits por segundo (bps) do sinal. A pergunta a se fazer é: qual a banda passante W necessária para se transmitir um sinal digital e 1/T bps? Ou, de forma inversa: quantos bits por segundos podemos transmitir em um meio físico cuja largura de banda é de W Hz?

7.3.1 Teorema de Nyquist

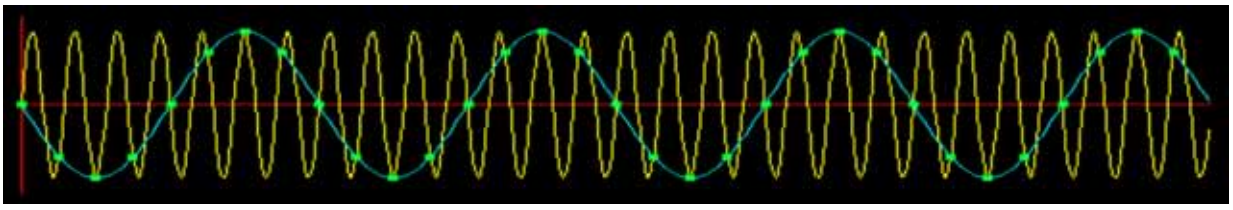
Em 1928, H. Nyquist formulou uma equação que define a taxa de transmissão máxima para um canal de banda passante limitada e imune a ruídos. Na prática este valor é inatingível mas serve de referência como um limite máximo teórico. Como veremos no decorrer da próxima seção, outras distorções podem ocorrer durante a transmissão de um sinal por um meio físico devido a fatores como atenuação, ruídos, etc. Alguns anos mais tarde, em 1948, Claude Shannon estendeu os resultados de Nyquist para o caso de um canal sujeito a ruído térmico.

Como inferir a taxa de transmissão máxima de um canal, dadas algumas características do canal em questão, como a sua banda passante e a razão sinal-ruído?

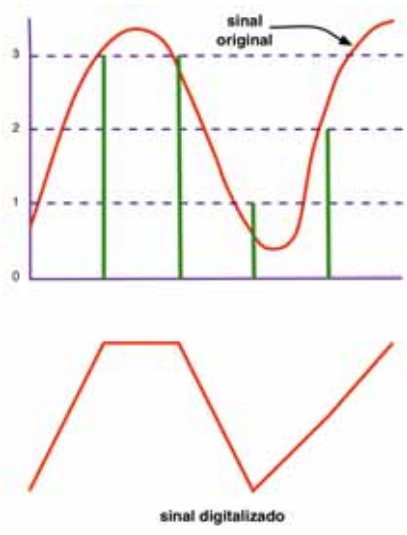
Inicialmente Nyquist formulou que dado um sinal arbitrário é possível coletar amostras desse sinal:



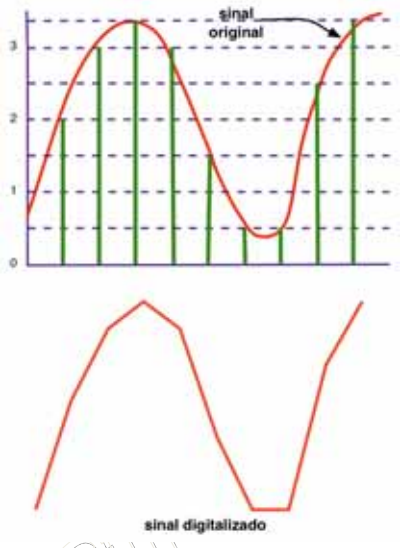
Nyquist provou que, se um sinal arbitrário é transmitido através de um canal de largura de banda W Hz, o sinal resultante da filtragem pode ser completamente reconstruído pelo receptor através da amostragem do sinal transmitido, a uma freqüência igual a no mínimo 2W vezes por segundo. Como exemplo, vejamos a análise de freqüência abaixo. O canal possui uma largura de banda W qualquer, o sinal arbitrário (original) corresponde a linha amarela, o sinal amostral (pontos coletados a um intervalo de 2W Hz) corresponde aos pontos em verde, e o sinal resultante, que corresponde a reconstrução do sinal original baseado apenas nos pontos de coleta, corresponde a linha em azul.



Nyquist demonstrou que esta é a freqüência mínima de amostragem necessária e, ao mesmo tempo, amostrar esse sinal a uma freqüência maior que 2W é inútil, já que as freqüências componentes que seriam recuperadas por uma tal amostragem já não existem no sinal devido à filtragem do canal. Quando tentamos recuperar um sinal arbitrário com taxas de amostragem diferentes do Teorema de Nyquist ocorre uma distorção do sinal original, a figura a seguir mostra o efeito da distorção sobre uma recuperação de sinal cujo intervalo das amostras coletadas foram inferiores a proposta de 2W de Nyquist.



Sinal Distorcido de uma Amostra menor que 2W



Sinal Resultante de uma Amostra de 2W

Para sinais digitais, isso corresponde a dizer que o número de transições de um nível de amplitude para outro no sinal original não pode ser maior do que 2W vezes por segundo. Em outras palavras, através de um canal de largura de banda igual a W Hz, pode-se transmitir um sinal digital de no máximo 2W bauds. Como 1 baud = $\log_2 L$ bps (onde L é o número de níveis utilizado na codificação), então a capacidade C do canal na ausência de ruído é dada por:

$$C = 2W\log_2 L \text{ bps}$$

Como exemplo, a tecnologia ADSL utiliza a banda passante de 4KHz à 2,2MHz, isso implica em uma largura de banda de aproximadamente 2,2MHz – 4KHz = 2,1MHz. Dada essa largura de banda W = 2.100.000Hz e supondo um número de níveis de codificação igual a 256 (2^8 = 256, padrão internacional para telefonia onde 8 corresponde a um código de 8 bits) temos que a capacidade do canal C, ou a velocidade máxima de bits por segundos no cana, é igual a: C = 2 * 2.100.00 * $\log_2$256 bps, que implica em C = 33.600.000 bps = 33,6Mbps. Entretanto, recordamos que esta equação de Nyquist só se aplica no aspecto teórico, visto que não releva aspectos como atenuações e ruídos na comunicação, o que na prática faz com que um link ADSL tenha uma capacidade máxima de canal em 24Mbps.

Como podemos observar, precisamos agora estudar as fontes de distorção dos sinais em transmissão para chegarmos nos valores exatos dos canais de comunicação que queremos contratar. Na próxima seção trataremos com exclusividade destas fontes, para em seguida aprendermos como calcular o canal com todos os elementos possíveis.

7.4 FONTES DE DISTORÇÃO DE SINAIS EM TRANSMISSÃO

Além dos efeitos de distorção dos sinais transmitidos oriundos da banda passante limitada do meio físico, outros fatores causarão distorções nos sinais durante a transmissão. Entre eles encontramos: os ruídos presentes durante a transmissão, a atenuação e os ecos. Passemos a analisar cada um desses fatores, seus principais efeitos e a forma de contorná-los.

7.4.1 Ruídos

Em qualquer transmissão, o sinal recebido consiste no sinal transmitido modificado por várias distorções impostas pelas características do meio físico adicionadas de outras distorções inseridas durante a transmissão devido à interferência de sinais indesejáveis denominados ruídos. O ruído é um dos mairoes limitantes do desempenho de sistemas de comunicação.

A quantidade de ruído presente numa transmissão é medida em termos da razão entre a potência do sinal e a potência do ruído, denominada razão sinal-ruído. Se representarmos a potência do sinal por S e a potência do ruído por N, a razão sinal-ruído é dada por S/N. É muito comum utilizar-se, ao invés desta razão diretamente, o valor $10\log_{10}(S/N)$. O resultado obtido é uma medida da razão sinal-ruído em uma unidade denominada *decibel* (dB). Uma razão de 10 corresponde a 10 dB; uma razão de 100 corresponde 20 dB; uma razão de 1.000 corresponde a 30 dB e assim por diante.

Ruídos podem ser classificados em quatro tipos: ruído térmico, ruído de intermodulação, crosstalk e ruído impulsivo.

O ruído térmico é provocado pela agitação dos elétrons nos condutores, estando, portanto, presente em todos os dispositivos eletrônicos e meios de transmissão.

O ruído térmico é uniformemente distribuído em todas as freqüências do espectro (sendo por isto freqüentemente citado como ruído branco) e sua quantidade é função da temperatura.

Quando sinais de diferentes freqüências compartilham um mesmo meio físico (através de multiplexação na freqüência – que veremos mais adiante) pode-se obter um ruído denominado de ruído de intermodulação. A intermodulação pode causar a produção de sinais em uma faixa de freqüências, que poderá perdurar a transmissão de outro sinal naquela mesma faixa. Este mau funcionamento acontece devido a defeitos em componentes do sistema ou devido a sinais com potência muito alta.

Crosstalk é um ruído bastante comum em sistemas telefônicos. Quem de nós ainda não teve a experiência de ser perturbado, durante uma conversação telefônica, por uma conversação travada por terceiros? É o fenômeno que comumente chamamos de “linha cruzada”. Este efeito é provocado por uma interferência indesejável entre condutores próximos que induzem sinais entre si.

Os tipos de ruídos descritos até aqui têm magnitudes e características previsíveis de forma que é possível projetar sistemas de comunicação que se ajuste a essas características. O ruído impulsivo, porém, é não contínuo e consiste em pulsos irregulares e com grandes amplitudes, sendo de prevenção difícil. Tais ruídos podem ser provocados por diversas fontes, incluindo distúrbios elétricos externos, falhas nos equipamentos etc.

O ruído impulsivo é, em geral, pouco danoso em uma transmissão analógica. Em transmissão de voz, por exemplo, pequenos intervalos onde o sinal é corrompido não chegam a prejudicar a inteligibilidade dos interlocutores. Na transmissão digital, o ruído impulsivo é a maior causa de erros de comunicação.

7.4.2 Lei de Shannon

Vinte anos depois de Nyquist, Claude Shannon, um grande físico e matemática, que entre suas obras está a teoria das comunicações, leis da criptografia, entre outros, provou também matematicamente, que um canal tem uma capacidade máxima limitada. A parte mais interessante de seu trabalho discute canais na presença de ruído térmico.

O principal resultado para as comunicações de dados de Shannon (conhecido como a Lei de Shannon) afirma que a capacidade máxima C de um canal (em bps) cuja largura de banda é W Hz, e cuja a razão sinal-ruído é S/N, é dada por:

$$C = W \log_2(1 + S/N)$$

Um canal de 3.000 Hz, por exemplo, com uma razão sinal-ruído de 30 dB (parâmetros típicos de uma linha telefônica) não poderá, em hipótese alguma, transmitir a uma taxa maior do que 30.000 bps, não importando quantos níveis de sinal se utilizem ou qual a freqüência de sinalização. É importante notar que este é um limite máximo teórico, e que, na prática, é difícil até mesmo se aproximar deste valor. Muito embora vários esquemas tenham sido propostos, a lei de Shannon constitui-se em um limite máximo intransponível.

7.4.3 Atenuação e Ecos

A potência de um sinal cai com a distância, em qualquer meio físico. Essa queda, ou atenuação, é, em geral, logarítmica e por isso é geralmente expressa em um número constante de decibéis por unidade de comprimento. A atenuação se dá devido a perdas de energia por calor e por radiação. Em ambos os casos, quanto maiores às freqüências transmitidas, maiores as perdas. A distorção por atenuação é um problema facilmente contornado em transmissão digital através da colocação de repetidores que podem regenerar totalmente o sinal original, desde que a atenuação não ultrapasse um determinado valor máximo. Para tanto, o espaçamento dos repetidores não deve exceder um determinado limite, que varia de acordo com a característica de atenuação do meio físico utilizado.

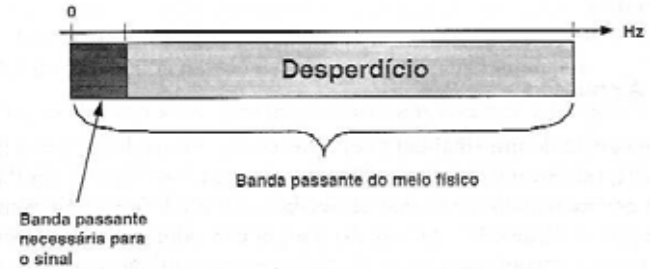
Ecos em linhas de transmissão causam efeitos similares ao ruído. Toda vez que há uma mudança de impedância numa linha, sinais serão refletidos e voltarão por esta linha, podendo corromper os sinais que estão sendo transmitidos.

Precauções para que a impedância de uma linha de transmissão não seja alterada podem ser tomadas para evitar a reflexão dos sinais. A utilização de terminadores e transceptores de alta impedância em redes em barra foram exemplificados no início do curso.

Em sistemas telefônicos, os ecos podem ser bastante desagradáveis quando percebidos em intervalos maiores que dezenas de milissegundos. Nesses sistemas é comum a utilização de canceladores de eco nos pontos onde é inevitável a alteração de impedância.

7.5 MULTIPLEXAÇÃO E MODULAÇÃO

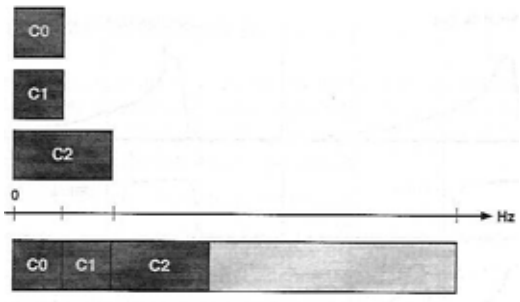
Sempre que a banda passante de um meio físico for maior ou igual à banda passante necessária para um sinal, podemos utilizar este meio para a transmissão do sinal. Na prática, a banda passante necessária para um sinal é, em geral, bem menor do que a banda passante dos meios físicos disponíveis, como mostra a figura abaixo:



A pergunta natural a se fazer neste momento é: não seria possível aproveitar a banda passante extra disponível para a transmissão de outros sinais? Mas especificamente, dados, por exemplo, três sinais (C_0, C_1, C_2) com a banda passante necessária indicada na figura, não seria possível transmiti-los simultaneamente através de um mesmo meio físico como apontado nesta mesma figura? A resposta a essa pergunta é sim, e a técnica que permite a transmissão de mais de um sinal em um mesmo meio físico é denominada multiplexação. Existem duas formas básicas de multiplexação: a multiplexação na freqüência (Frequency Division Multiplexing – FDM) e a multiplexação no tempo (Time Division Multiplexing – TDM).

7.5.1 Multiplexação na Frequência (FDM)

Em primeiro lugar, se passarmos um filtro em cada um dos sinais da imagem abaixo



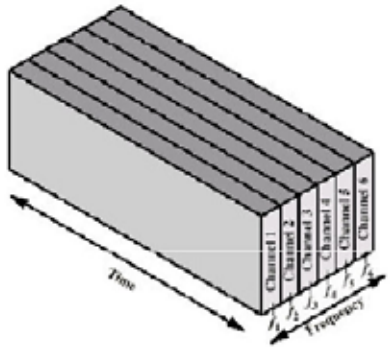
de forma a preservar somente a faixa relativa à banda passante necessária de cada um deles, teremos dado o primeiro passo para alojar esses três sinais na forma desejada, sem que um sinal interfira no outro. O passo seguinte é deslocar a faixa de freqüência original do segundo e do terceiro sinal de forma que eles passem a ocupar as três faixas disjuntas, sem sobreposição.

Felizmente, técnicas que permitem esse deslocamento ou *shift* de freqüências são conhecidas e denominadas técnicas de modulação. Dessa forma, os três sinais podem ser transmitidos no meio físico, cada um deles ocupando uma banda ou canal distinto com tamanho necessário para a sua transmissão. Como os sinais foram previamente filtrados de acordo com a sua banda passante necessária, a informação de cada um deles está preservada e contida naquela faixa de freqüências na qual está sendo transmitido e em nenhuma outra.

Equipamentos capazes de realizar modulação e demodulação de sinais são denominados MODEM (moduladores/demoduladores). Veremos agora essas técnicas de modulação.

7.5.2 Técnicas de Modulação

A introdução dos sistemas de transmissão digital utilizando a tecnologia de Modulação no início da década de 1970, revolucionou os sistemas de telecomunicações impulsionando ainda mais o processo de reestruturação geral que elevou o nível de competitividade que passou a caracterizar os mercados de produtos e serviços.



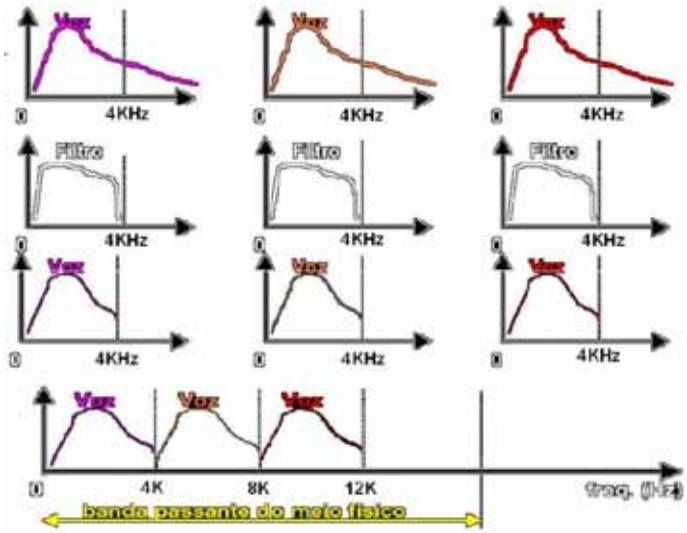
Até a introdução da tecnologia de modulação por pulsos, os sistemas eram conhecidos como AM-DSB (Amplitude Modulation – Double Side Band), AM-DSB/SC (Supried Carrier – Portadora Suprimida) e AM-SSB (Amplitude Modulation – Single Signal Band), bem como sistemas de FM (Frequency Modulation) e outros. Tais sistemas operam continuamente, ou seja, durante todo o tempo do sinal a ser transmitido.

Para facilitar a transmissão do sinal através dos meios físicos e adequar as freqüências aos sistemas de comunicação se utiliza o que chamamos de onda portadora, sobre a qual é transmitido o sinal.

A onda portadora é um sinal senoidal caracterizado por três variáveis: amplitude, freqüência e fase. A amplitude é a medida da altura da onda para voltagem positiva ou para voltagem negativa. Também definida como crista da onda, a amplitude do sinal digital é igual à diferença da voltagem para o degrau entre 0 e 1. Iniciando na voltagem zero, a onda cresce atinge a amplitude, decresce, se anula, atinge sua amplitude negativa e volta a crescer até se anular novamente. Essa seqüência compõe um ciclo.

Modulação é o processo na qual a informação é adicionada a ondas eletromagnéticas. É assim que qualquer tipo de informação, até a voz humana ou transação de dados numa aplicação interativa é transmitida numa onda eletromagnética. O transmissor adiciona a informação numa onda básica de tal forma que poderá ser recuperada na outra parte através de um processo reverso chamado demodulação.

A imagem ao lado destaca a técnica de multiplexação por freqüência, onde existe um único meio físico para transmissão e cuja a banda passante é muito superior a largura de banda necessária para a aplicação de voz, que geralmente necessita de 4KHz apenas. Dessa forma, através da modulação é possível organizar o espaço dentro da banda passante do meio físico de forma a comportar sem sobreposições três aplicações de voz simultâneas.



Para este exemplo podemos dizer que, existem três canais independentes, ou que existem três linhas telefônicas independentes, passando pelo mesmo meio físico e que somente filtros ou modems sabem como interpretar tais canais.

Um processo de modulação consiste em modificar o formato da informação elétrica com o objetivo de transmiti-la com a menor potência possível, com a menor distorção possível, facilidade de recuperação da informação original e ao menor custo possível.

Nas modernas redes de telecomunicação, a informação é transmitida, transformando em uma das duas características da onda: a amplitude e a freqüência.

Modulação em Amplitude - AM (Amplitude Modulation) - usa o sistema de chaveamento de amplitude ASK (Amplitude Shift Keying). É usada na comunicação de voz, na maioria das transmissões de LAN's, mas pouco indicada para WLAN porque é muito sensível ao ruído;

Modulação por freqüência - FM (Frequency Modulation) - usa o chaveamento de freqüência FSK (Frequency Shift Keying).

Pela modulação caracterizamos a forma de apresentar a informação que se transforma em tráfego. Podemos ter modulação analógica e digital:

Modulação analógica: Também classificada como modulação de onda continua (CW), na qual a portadora é uma onda consenoidal e o sinal modulante é um sinal analógico ou contínuo;

Modulação digital: Também denominada modulação discreta ou codificada. Utilizada em casos que se está interessado em transmitir uma forma de onda ou mensagem que faz parte de um conjunto finito de valores discretos representando um código.

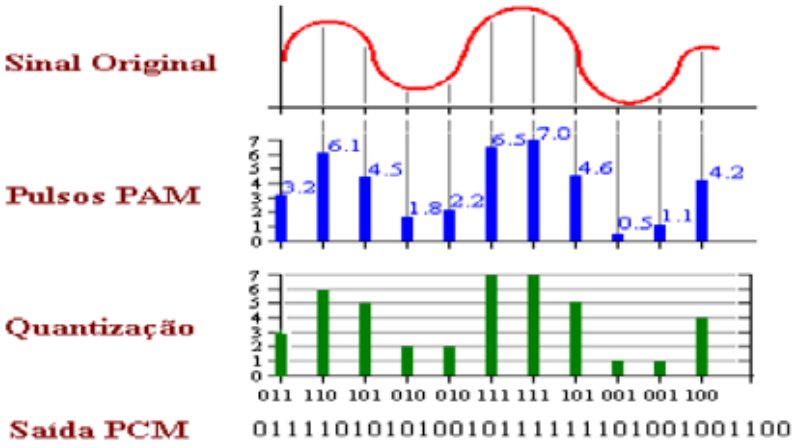
Os sistemas baseados em sinal e a modulação digital oferecem grandes vantagens sobre os sistemas analógicos, por exemplo, alta fidelidade, independência do tempo e da fonte dos sinais que podem ser codificados.

Uma desvantagem está no elevado custo dos equipamentos de transmissão, principalmente para aplicações em tempo real, pois são precisos complexos e caros circuitos para que a comunicação digital possa ser realizada em tempo real.

A modulação digital tem preferência sobre a analógica devido a um fator fundamental: a informação transmitida na forma digital pode ser regenerada, replicada e retransmitida, mantendo-se livre de distorções. Esta vantagem, entretanto, possui um certo custo: o sinal modulado digitalmente ocupa maior largura de faixa que seu correspondente modulado analogicamente. Outra vantagem da modulação digital consiste na possibilidade de multiplexação de sinais de informação originalmente analógica juntamente com dados provenientes de computadores os quais já são digitais por natureza.

Enquanto que a modulação analógica está diretamente associada a equipamentos MODEM, a modulação digital está associada a equipamentos CODECs.

Um padrão de modulação digital é o PCM (Pulse Code Modulation), que converte o sinal analógico em pulsos por amplitude (PAM) e quantizam os PAMs aproximando seus valores em um inteiro de bits.



O PCM é visto hoje como um padrão de CODEC para as transmissões digitais. O CODEC (Coder/Decoder) são equipamentos, hardwares e/ou softwares, responsáveis por converter os sinais analógicos em digital. Um exemplo seria a transmissão das linhas telefônicas digitais, o PCM é um, entre vários outros CODECS, capaz de modular o sinal analógico para digital, para este codec a banda passante necessária para a voz é de 4.000Hz, um padrão internacional já homologado, e adota uma taxa de amostragem de Nyquist de 8.000 amostras por segundo, um valor também já homologado internacionalmente para as comunicações digitais de voz. Dessa forma, o CODEC PCM (padrão para linhas digitais de voz) corresponde a uma largura de banda fixa de 8.000 x 8 (bits) = 64Kbps

Encontramos freqüentemente o PCM nos ambientes de VoIP (Voice over IP), onde o PCM é visto como a opção padrão para o tratamento de voz pela Internet, o que significa dizer que os roteadores e os enlaces entre a estação de origem e destino precisam ter, pelo menos, e constantemente, 64Kbps. Menos do que isso o CODEC finaliza a transmissão digital.

Quando realizamos uma comparação entre o PCM e outros CODEC estamos analisando o algoritmo de compactação do sinal sobre o meio digital, ou seja, algoritmos mais sofisticados de CODEC possibilitam uma qualidade excepcional de recuperação de sinal com uma menor taxa de amostragem de Nyquist, dessa forma é possível realizar a mesma transmissão com uma largura de banda menor.

7.5.3 Multiplexação no Tempo (TDM)

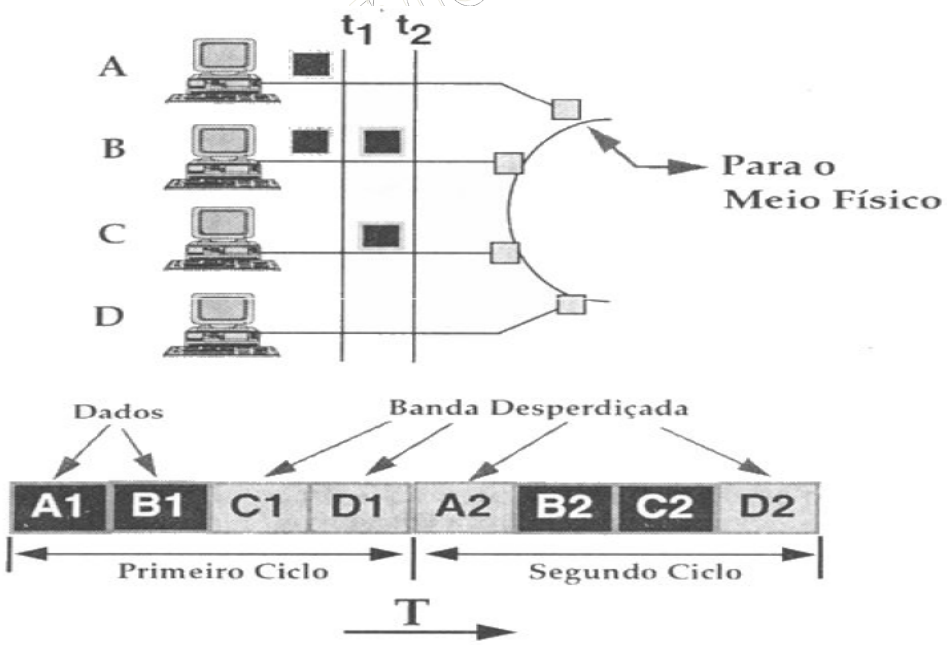
TDM é uma tecnologia digital, fazendo a multiplexação de sinais digitais. Conforme já mencionado, podemos compartilhar um meio físico por varias estações não só pela multiplexação na freqüência, mas também pela multiplexação no tempo. A multiplexação por divisão do tempo (Time Division Multiplexing – TDM) se beneficia do fato de que a capacidade (em quantidade de bits por segundo) do meio de transmissão, em muitos casos, excede a taxa média de geração de bits das estações conectadas ao meio físico. Quando isso ocorre, vários sinais podem ser transportados por um único caminho físico, intercalando-se porções de cada sinal no tempo.

O TDM é caracterizado pela presença de um elemento de sincronismo. O intervalo entre 2 padrões define um quadro (*frames*). Cada quadro é subdividido em fatias de tempos (*slots*). O tempo é dividido em intervalos regulares. Cada sub-canal tem direito a um intervalo de tempo para transmitir seus dados. Toda a banda do canal fica disponível. Veremos agora como compreender melhor essas informações técnicas sobre o TDM.

A multiplexação no tempo pode ser classificada em síncrona ou assíncrona.

- TDM Síncrono

  O tempo é dividido em frames de tamanho fixo, que por sua vez são divididos em intervalos de tamanho fixo, denominados canais. Por exemplo, ao dividirmos o tempo em 10 frames de tamanho fixo, e definirmos que a estação 1 utilizará o canal 1 destes 10 frames, implica em dizer que o transmissor do canal 1 só poderá transmitir dados no intervalo 1, devendo aguardar novamente sua vez após completado o ciclo de 10 frames. Havendo dados cujo tempo de transmissão excedam o tempo do seu intervalo, o mesmo deverá aguardar por uma nova rodada do ciclo para continuar a transmitir. Da mesma forma, caso os demais 9 canais não estejam sendo utilizados, mesmo assim a estação do canal 1 deverá aguardar todos os ciclos, com seus respectivos tempos por canais, para poder transmitir.

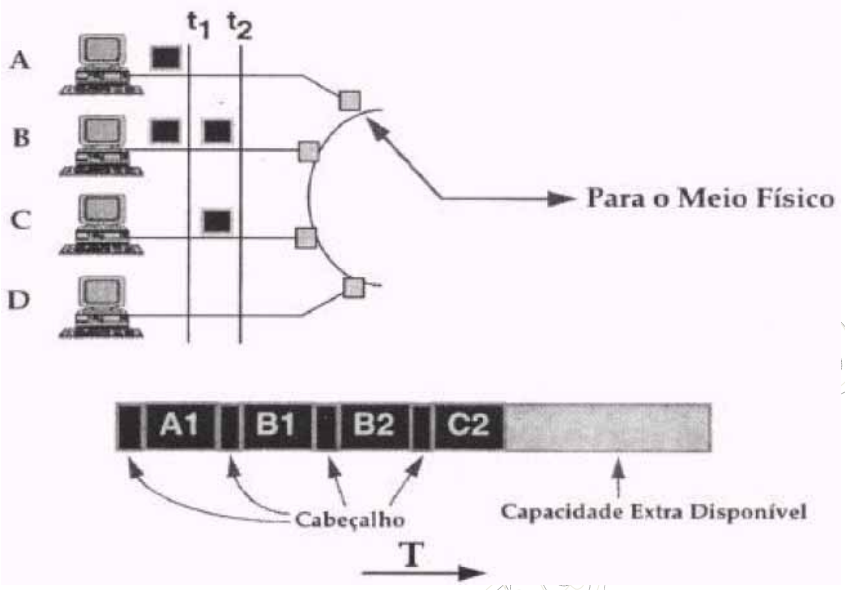


  Quando o canal não tem demanda regular e contínua, o TDM síncrono não é eficiente.

- TDM Assíncrono ou STDM (*Statistical Time Division Multiplexing*)

  Corresponde a uma variação do TDM Síncrono de forma a favorecer a eficiência em canais que não possuem demanda regular. Neste esquema não há alocação de canal para uma fonte. Uma fonte pode utilizar qualquer intervalo de tempo, desde que este não esteja sendo usado por outra conexão.

  Neste modelo as parcelas de tempo são alocadas dinamicamente de acordo com a demanda das estações. Nenhuma capacidade é desperdiçada, pois o tempo não usado fica disponível para outra fonte. Entretanto necessita que cada unidade de informação conste um cabeçalho, contendo o endereço de origem e destino, pois todos os dados de todas as estações ficam misturados no quadro a ser enviado.



Assim como o FDM, o TDM é utilizado sobre cabos coaxiais, microondas e fibras óticas. Infelizmente só pode ser usado para transmissões digitais, porém é possível converter sinais analógicos para digitais a fim de aproveitar os benefícios do TDM, como é o caso dos entroncamentos entre centrais telefônicas, utilizando o mecanismo de CODECs.

Além dos dois tipos básicos ainda é possível realizar uma combinação entre, formando redes híbridas ou canais multiponto em redes de banda larga.

7.5.4 Técnicas de Transmissão

Conforme vimos anteriormente existem algumas técnicas para transmitir em um mesmo meio físico sinais diferentes, em resumo é possível realizar o seguinte tipo de combinação:

- Dado Digital – Sinal Digital (redes RDSI como a Manchester)
- Dado Digital – Sinal Analógico (codecs como o PCM)
- Dado Analógico – Sinal Digital (modulação)
- Dado Analógico – Sinal Analógico (as tradicionais redes de telefonia analógicas)

As técnicas de modulação vistas condicionam o surgimento de dois tipos de transmissão:

- Sinalização em banda básica (baseband) (adotada em linhas digitais)
    - Não adotam modulações e sim codecs;
    - Toda a frequência do meio é utilizada para o sinal (uni-canal);
    - Possibilitam transmissões em alta velocidade;
    - Não devem ser utilizados em canais susceptíveis a ruídos;
    - Ideal para redes locais;
    - Exemplo: Ethernet, Wireless, Bluethooh, TV, etc.

- Sinalização em banda larga (broadband) (adotada em linhas analógicas)
    - Adota a técnica de modulação por freqüência - FDM (multi-canais);
    - Topologia exclusivamente em barra;
    - Os sinais são propagados em modo simplex (em função do custo);
    - Uso de dois cabos (um para transmitir – TX ou upstreamer e outro para receber – RX ou downstreamer)
    - Exemplo: ADSL, Cabo, Satélite, LPDC
    - Necessidade de dois caminhos e uma central repetidora:

caminho da transmissão (inbound) -> headend -> caminho de retorno (outbound)

7.6 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. Seja a rede uma LAN, MAN ou WAN, existem sempre recursos compartilhados. Nas LANs e MANs, por exemplo, a utilização de topologias com meio compartilhado do tipo barra é bastante comum. Nas WANs, a utilização da topologia parcialmente ligada fará com que os caminhos entre pares de estações tenham que utilizar, muitas vezes, os mesmos enlaces, o que determina o compartilhamento desses enlaces durante o funcionamento da rede. A alocação desses recursos está, como podemos perceber, intimamente ligada à forma de multiplexação dos meios de transmissão.

As principais formas de comutação são denominadas comutação de circuitos e comutação de pacotes. Passamos agora a analisar essas formas de comutação juntamente com variações que têm surgido para aumentar o desempenho desses esquemas.

7.6.1 Comutação de Circuitos

A comunicação via comutação de circuitos pressupõe a existência de um caminho dedicado de comunicação entre duas estações. A comunicação via comutação de circuitos envolve três fases:

1. Estabelecimento do circuito: um circuito fim a fim é estabelecido, determinando e alocando uma rota entre as estações, onde, em cada enlace, um canal é alocado e permanece dedicado a essa conexão até a hora da desconexão do circuito.
2. Transferência de informação: uma vez estabelecida a conexão, os dados podem ser transmitidos e recebidos pelas estações envolvidas como se não houvessem intermediários no circuito;

3. Desconexão do circuito: após um certo período de tempo a conexão pode ser encerrada, em geral pela ação de uma das estações envolvidas. Sinais de controle devem ser propagados por todos os nós intermediários do circuito de forma que todos os caminhos sejam desalocados.

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. Isso significa que, caso o tráfego entre as estações não seja constante e contínuo, a capacidade do meio físico será desperdiçada. Em compensação, existe a garantia de que uma taxa de transmissão está sempre disponível quando as estações desejam se comunicar, pois não há contenção alguma de recursos.

O caminho dedicado entre a origem e o destino pode ser:

- Um caminho físico formado por uma sucessão de enlaces físicos que permanecem alocados a conexão até o momento da desconexão.
- Uma sucessão de canais de freqüência alocados em cada enlace. Cada nó intermediário associa um canal de freqüência de um enlace a um canal de freqüência de um outro enlace e forma que, ao receber um sinal de uma porta em uma determinada freqüência, filtra, demodula, remodula na outra freqüência, e transmite o sinal na porta de saída no canal associado.
- Uma sucessão de canais de tempo alocados em cada enlace. Cada nó intermediário associa um canal TDM (síncrono) de uma linha a uma canal TDM (síncrono) em outra linha. Cada um desses nós desmultiplexa e torna a multiplexar os sinais de algumas portas em outras portas para fechar os circuitos desejados.

Um exemplo de comutação de circuito está nos PBXs, ou centrais telefônicas privadas.

7.6.2 Comutação de Mensagens

Na comutação de mensagens não é necessário o estabelecimento de um caminho dedicado entre as estações. Ao invés disso, se uma estação deseja transmitir uma mensagem (uma unidade lógica de informação), ela adiciona o endereço de destino a essa mensagem que será então transmitida pela rede de nó em nó. Em cada nó, a mensagem inteira é recebida e o próximo caminho da rota é determinado com base no endereço contido na mensagem.

Esse caminho pode se encontrar ocupado pela transmissão de outra mensagem e, ainda, outras mensagens já podem estar esperando para serem transmitidas por esse mesmo caminho. Nesse caso, a mensagem espera numa fila até que chegue a sua vez de ser transmitida e o caminho esteja liberado, quando então a transmissão se inicia. Assim, uma mensagem caminho de nó em nó pela rede utilizando apenas um canal por vez, sendo armazenada e retransmitida em cada nó (processo conhecido como *store-and-forward*).

Pode-se observar algumas características na comutação de mensagens em relação à comutação de circuitos:

- O aproveitamento das linhas de comunicação é maior, já que os canais podem ser compartilhados por várias mensagens ao longo do tempo, devido ao fato de não haver alocação dos canais; mensagens são transmitidas por demanda.
- Quando o tráfego se torna alto em uma rede de comutação de circuitos, pedidos de novas conexões podem ser recusados devido à falta de recursos ou caminhos livres. As mensagens são sempre aceitas em uma rede comutação de mensagens, o tempo de transferência é que aumento devido às filas que as mensagens encontrarão em cada nó de comutação da rede.

7.6.3 Comutação de Pacotes

A comutação de pacotes é semelhante à comutação de mensagens. A principal diferença está no fato de que o tamanho da unidade de dados transmitida na comutação de pacotes é limitado. Mensagens com tamanho acima de um limite devem ser quebradas em unidades menores denominadas pacotes. Pacotes de uma mesma mensagem podem estar em transmissão simultaneamente pela rede em diferentes enlaces, o que reduz o atraso de transmissão total de uma mensagem. Além disso, redes com comutação de pacotes requerem nós de comutação com menor capacidade de armazenamento e os procedimentos de recuperação de erros para pacotes são mais eficientes do que para mensagens.

A técnica de comutação de pacotes é também uma técnica store-and-forward, dado que pacotes caminham de nó em nó pela rede, sendo armazenados e retransmitidos sucessivamente. Em cada nó, um pacote inteiro é recebido e o próximo caminho da rota é escolhido. Logo, vemos que cada pacote deve conter a informação de seu destino (endereço de destino) de forma a possibilitar o roteamento correto.

7.7 TÉCNICAS DE DETECÇÃO DE ERROS

Apresentamos no início desta lição vários fenômenos que podem causar erros de transmissão, como os ruídos. Na impossibilidade de eliminar totalmente esses fenômenos, sistemas de comunicação devem ser projetados de forma a possibilitar a recuperação da informação perdida. O primeiro passo para qualquer esquema de tratamento de erros é a sua detecção. Reconhecer que um quadro foi recebido com erro irá permitir que se tomem as providencias necessárias, que poderão variar de acordo com as necessidades das aplicações e com as características dos dados transmitidos. Nesta seção abordaremos apenas as técnicas para a sua detecção.

Todos os métodos de detecção de erros são baseados na inserção de bits extras na informação transmitida. Esses bits consistem em informação redundante, isto é, que pode ser obtida a partir da informação original.

Esses bits são computados pelo transmissor através de algum algoritmo que tem como entrada os bits originais a serem transmitidos. Após computar esses bits, o transmissor os acrescenta aos bits de informação propriamente dita, para então prosseguir com a transmissão do quadro. Quando o quadro é recebido, o receptor, conhecendo o algoritmo utilizado pelo transmissor, pode recomputar os bits de redundância e compará-los com os respectivos bits recebidos no quadro. Se eles forem diferentes, detectou-se a presença de um erro.

Vários algoritmos para a geração de bits de redundância já foram propostos e podem ser encontrados na literatura, como Stallings e Tanembaum. Iremos apresentar resumidamente duas das principais técnicas conhecidas pelos nomes de paridade e CRC (*Cyclic Redundancy Checks*).

7.7.1 Paridade

A verificação de paridade é um dos mecanismos mais simples para detecção de erros (“parity check”): a cada caractere transmitido é acrescentado um bit de tal modo que o total de bits 1 seja par (“even parity”) ou ímpar (“odd parity”). É habitual a utilização de paridade par para comunicações assíncronas e a paridade ímpar para comunicações síncronas.

Esta técnica consiste em acrescentar um bit extra ao caractere, isto é, emprega a técnica de paridade que pode ser paridade par ou paridade ímpar, ou seja, a soma dos bits ligados (1) de um caracteres deve ser igual a um valor ímpar ou par.

Além dos oito bits de caractere que são gerados, a estação transmissora adiciona um bit de paridade para cada caractere e a soma desses nove bits deverá manter-se sempre ímpar ou par, dependendo da técnica de paridade empregada. Não há restrição ao uso da técnica de paridade em relação ao código utilizado pelo equipamento (Baudot, ASCII, EBCDIC, etc.). Exemplos:

Paridade Par:

Carácter	Bit de Paridade	Sequência a Transmitir
1000100	0	10001000
1110000	1	11100001

Paridade Ímpar:

Carácter	Bit de Paridade	Sequência a Transmitir
1000100	1	10001001
1110000	0	11100000

O equipamento transmissor calcula o bit de paridade para cada caractere transmitido. O receptor calcula um novo bit de paridade em cima dos bits recebidos e compara este bit com aquele enviado pelo transmissor. Se forem iguais, a transmissão é considerada correta, se não, haverá necessidade de retransmissão do caractere. Caso haja um número par de bits com erro, a técnica não consegue detectar, pois a verificação de bits "1"s do caractere recebido permanecerá par ou ímpar, de acordo com o método, satisfazendo ao bit de paridade. Entretanto, a prática mostra que a maioria dos erros é simples.

7.7.2 CRC

O método CRC (Cyclic Redundancy Checking), embora use uma técnica mais complexa, é bem mais eficiente que os anteriores.

A técnica de verificação cíclica é executada por ambas as estações, transmissora e receptora, e consiste na divisão de todos os bits de um bloco por um valor binário constante (polinômio gerador). O quociente é desprezado e o resto desta operação será o caractere de verificação que será transmitido.

O CRC, também conhecido como método de detecção polinomial, é um processo de verificação de erros mais sofisticado que os anteriores, permitindo que se detecte praticamente a ocorrência de qualquer grupo de erros.

Alguns polinômios geradores são largamente utilizados e padronizados. Como exemplo, temos:

	Polinômio gerador	**Comprimento de caractere**	**Capacidade de detecção de erros**
CRC-12	$x^{12}+x^{11}+x^{3}+x^{2}+x+1$	6 bits	Até 12 erros simultâneos
CRC-ITU	$x^{16}+x^{12}+x^{5}+1$	8 bits	Até 16 erros simult. ou 99% dos casos
CRC-16	$x^{16}+x^{15}+x^{2}+1$	8 bits	Até 16 erros simult. ou 99% dos casos
CRC-32	$x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^{8}+x^{7}+x^{5}+x^{4}+x^{2}+x+1x$	8 bits	Até 32 erros simult. ou 99% dos casos

O esquema baseado em CRC-12 é utilizado em seqüências de caracteres de seis bits gerando um FCS de 12 bits. Tanto CRC-16 quanto CRC-ITU são populares para seqüencias de caracteres de oito bits, na Europa e EUA, respectivamente, ambos resultando em FCS de 16 bits. O CRC-32 foi o escolhido pelo comitê IEEE-802 para ser utilizado em redes locais, gerando um FCS de 32 bits. Vejamos como funcionam essas técnicas:

- *Na transmissão:*
  1. Os dados de informação a serem transmitidos são transformados num polinômio D(x), em função dos "0"s e "1"s.;
  2. O polinômio D(x) é multiplicado pelo termo de maior grau de um polinômio gerador G(x). O resultado desta multiplicação será um novo polinômio D'(x);
  3. Fazemos a divisão do polinômio D'(x) por G(x);
  4. O resto desta divisão R(x) será adicionado no fim da transmissão de D(x). Dependendo do tipo de protocolo utilizado, esse "resto" leva diferentes nomes (BCC no protocolo BSC e FCS no protocolo X.25).

- *Na recepção:*
  1. Os dados recebidos serão divididos pelo mesmo polinômio gerador G(x);
  2. Se o resto desta divisão for igual a zero, significa que não houve erros na transmissão; caso contrário, foi detectado erro na transmissão, sendo necessária a retransmissão da informação enviada anteriormente.

Exemplo - dados a seqüência e o polinômio gerador, verificar erros de transmissão:
D(x)=10110110 (seqüência a ser transmitida)
G(x)= $x^3+ x^2+x^1$ (polinômio gerador simplificado)

- Transmissão
  1. Definição de D(x):

$$D(x)= 1x^7+0x^6+1x^5+1x^4+1x^3+0x^2+1x^1+1x^0+1$$

  2. Definição de D'(x), multiplicando-se D(x) pelo termo de maior grau de G(x), no caso x^3, tem-se:

$$x^7+x^5+x^4+x^3+x+1$$
$$\times\ x^3$$
$$D'(x)=x^{10}+x^8+x^7+x^6+x^4+x^3$$

  3. Definição de R(x), dividindo-se D'(x) por G(x):

$$x^{10}+x^8+x^7+x^6+x^4+x^3 \ /\ x^3+x^2+x$$
$$x^{10}+x^9+x^8 \qquad\qquad x^7+x^6+x^5+x^4+x^3+1$$
$$x^9+x^7+x^6$$
$$x^9+x^8+x^7$$
$$x^8+x^6+x^4$$
$$x^8+x^7+x^6$$
$$x^7+x^4+x^3$$
$$x^7+x^6+x^5$$
$$x^6+x^5+x^4+x^3$$
$$x^6+x^5+x^4$$
$$x^3$$
$$x^3+x^2+x$$
$$\boxed{R(x)= x^2+x}$$

  4. R(x)=x^2+x=110 (em binário) que será enviado ao final da sequência D(x), ou seja: 10111011110.

- Recepção

1. Converter a seqüência recebida em polinômio:

$$1x^{10}+0x^{9}+1x^{8}+1x^{7}+1x^{6}+0x^{5}+1x^{4}+1x^{3}+1x^{2}+1x^{1}+0x^{0}$$

2. Dividir a seqüência recebida pelo mesmo polinômio gerador G(x):

$$x^{10}+x^{8}+x^{7}+x^{6}+x^{4}+x^{3}+x^{2}+x \;/\; x^{3}+x^{2}+x$$
$$x^{10}+x^{9}+x^{8} \qquad\qquad x^{7}+x^{6}+x^{5}+x^{4}+x^{3}+1$$
$$x^{9}+x^{7}+x^{6}$$
$$x^{9}+x^{8}+x^{7}$$
$$x^{8}+x^{6}+x^{4}$$
$$x^{8}+x^{7}+x^{6}$$
$$x^{7}+x^{4}+x^{3}$$
$$x^{7}+x^{6}+x^{5}$$
$$x^{6}+x^{5}+x^{4}+x^{3}$$
$$x^{6}+x^{5}+x^{4}$$
$$x^{3}+x^{2}+x$$
$$x^{3}+x^{2}+x$$
$$0$$

O resto=0 indica que não houve erros na transmissão, caso contrário seria necessário retransmitir toda a seqüência.

7.8 MEIOS FÍSICOS DE TRANSMISSÃO

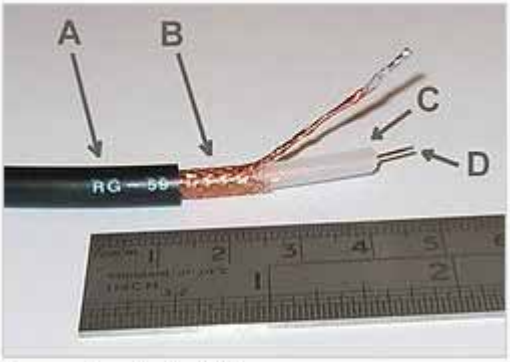
Os meios físicos de transmissão diferem entre si com relação a sua banda passante, potencial para conexão ponto-a-ponto ou multiponto, limitação geográfica em função da atenuação característica do meio, imunidade à ruído, custo, disponibilidade, confiabilidade, entre outros parâmetros de comparações de redes já visto no curso.

A escolha do meio de transmissão adequado às aplicações é extremamente importante não só pelos motivos mencionados, mas também pelo fato de que ele influencia diretamente no custo dos projetos de rede.

Qualquer meio físico capaz de transportar informações eletromagnéticas é passível de ser usado em redes de computadores. Atualmente o mercado dispõe de tecnologias baseadas em cabos (*wired*), e tecnologias móveis que dispensam o uso de cabos (*wireless*). Os mais comumente utilizados nas redes cabeadas são o par trançado, o cabo coaxial e a fibra ótica. Nas redes sem fio os mais comumente utilizados são a radiodifusão, infravermelho, enlaces de satélite e microondas. Vejamos suas características.

7.8.1 Cabo Coaxial

O cabo coaxial foi um dos primeiros que surgiu para as ligações em redes de computadores. Desde a década de 70 até 2010, ainda é muito utilizado por diversas aplicações, como: antena coletiva de televisão, antena de televisão por assinatura, redes de computadores, equipamentos de comunicação de dados como o E1, entre outros.



A: revestimento de plástico
B: tela de cobre
C: isolador dialétrico interno
D: núcleo de cobre

É um tipo de cabo condutor usado para transmitir sinais. Este tipo de cabo é constituído por diversas camadas concêntricas de condutores e isolantes, daí o nome coaxial. É constituído por um fio de cobre condutor revestido por um material isolante e rodeado duma blindagem. Este meio permite transmissões até frequências muito elevadas e para longas distâncias.

O cabo coaxial possui como vantagem sua blindagem adicional, que o protege contra o fenômeno da indução, causado por interferências elétricas ou magnéticas externas. Essa blindagem constitui-se de uma malha metálica (condutor externo) que envolve um condutor interno isolado

A principal razão da sua utilização deve-se ao fato de poder reduzir os efeitos e sinais externos sobre os sinais a transmitir, por fenômenos de IEM (Interferência Electromagnética).

Os cabos coaxiais geralmente são usados em múltiplas aplicações desde áudio ate as linhas de transmissão de freqüências da ordem dos gigahertz . A velocidade de transmissão é bastante elevada, na ordem de megabits por segundo, sem necessidade de regeneração do sinal e sem distorções ou ecos, devido a tolerância aos ruídos, graças à malha de proteção desses cabos. Adotam a topologia física em barramento com cabos dispostos em série (Ethernet) ou estrela (ARCNet). E necessitam de aterramento para o cabo e terminadores a fim de evitar interferências.

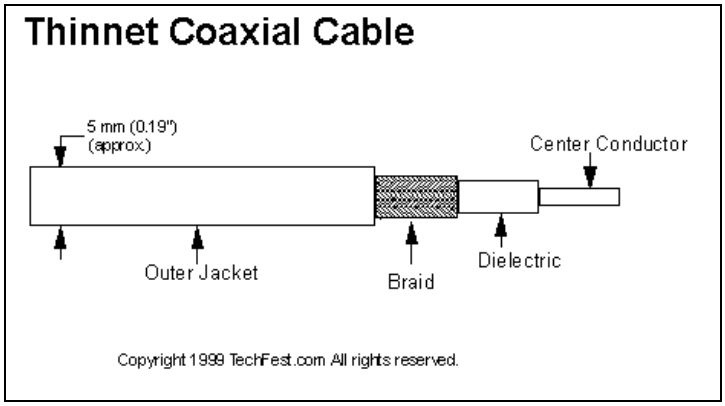
Existe uma grande variedade de cabos coaxiais, cada um com características específicas. Alguns são melhores para transmissao em alta frequência, outros têm atenuação mais baixa, oturos são mais imunes a ruídos e interferências, etc. Os cabos de mais alta qualidade não são maleáveis e são difícies de instalar, mas cabos de baixa qualidade podem ser inadequados para altas velocidades e longas distâncias.

CABO	(ohms)	DESCRIÇÃO
RG-6/U	75	Cable Modem, Televisão a Cabo, Televisão a Satélite
RG-8 e RG-11	50	Thicknet - Cabo Grosso para Redes de Computadores
RG-58/U	50	Thinnet – Cabo Fino para Redse de Computadores
RG-60/U	50	Televisão de Alta Definição e Internet a Cabo de Alta Velocidade
RG-174/U	50	Pigtails em Redes Wireless

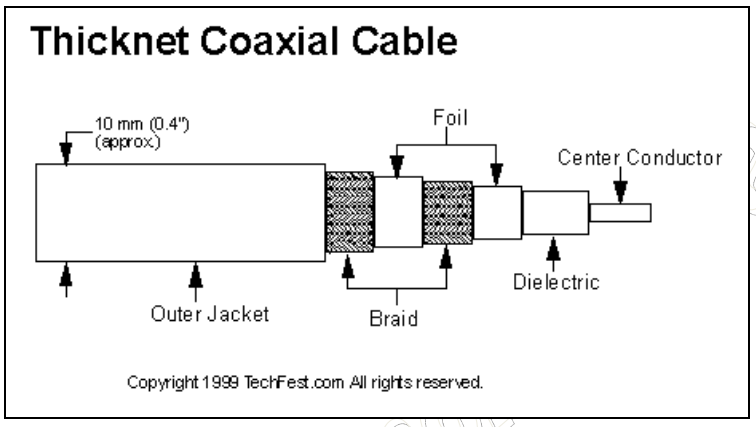
A maioria dos sistemas com transmissão em banda básica utiliza o cabo com impedância característica de 50 ohms, ao invés do cabo de 75 ohms comumente utilizado nas TVs a cabo e nas redes em banda larga. Isso se deve ao fato de que na transmissão em banda básica o cabo de 50 ohms sofre menos reflexões devido às capacitâncias introduzidas na ligação das estações ao cabo, além de possuir uma maior imunidade a ruídos eletromagnéticos de baixa freqüência. Sistemas com transmissão em banda larga utilizam a tecnologia desenvolvida para os componentes CATV (Community Antenna Television), incluindo o cabo coaxial de 75 ohms.

Para as redes de computadores os tipos mais utilizados são apenas 02:

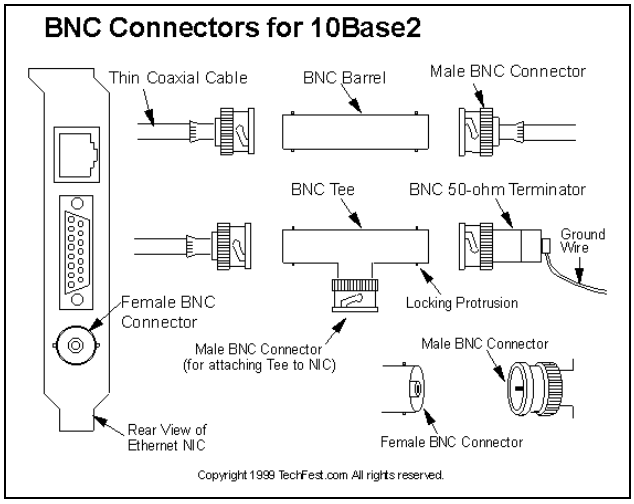
- Coaxial Fino (Thinnet / 10Base2): mais leve, flexível, barato e fácil de instalar que o Coaxial Grosso. Utiliza o tipo de cabo RG-58 com impedância de 50 omhs. Suporta um máximo de 30 nós e alcança velocidades entre 2,5Mbps à 10Mbps, podendo ter comprimento de cabo máximo de 185 metros por segmento.



- Coaxial Grosso (Thicknet / 10Base5): usa o dobro do diâmetro do Thinnet (12mm), adotando o cabo RG-11 com impedância de 50 ohms. Também conhecido como Cabo Ethernet Padrão, e é utilizado como backbone para interconexão de LANs Thinnet. Suporte um máximo de 100 nós e também pode alcançar velocidades entre 2,5Mbps à 10Mbps, podendo ter comprimento de cabo máximo de 500 metros por segmento.

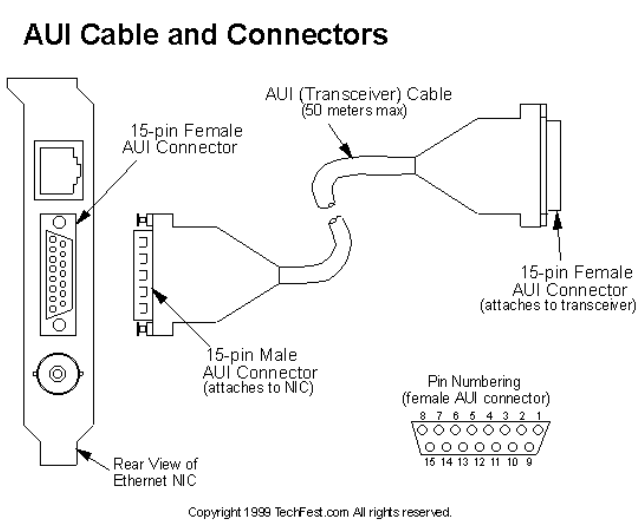


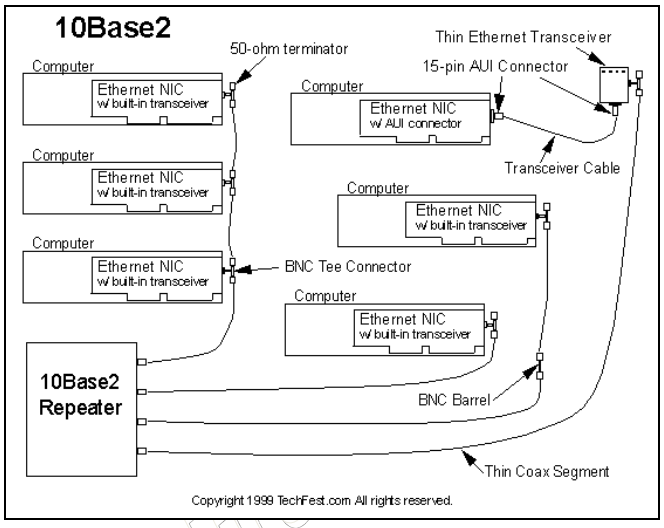
As primeiras redes de computadores que adotavam o padrão de cabo coaxial, adotavam interfaces de rede específicas:



As placas de redes eram ligadas a um conector especial, chamado conector BNC-T, que funcionava para ligar o computador ao barramento e outra estação de trabalho. Em seguida utilizava-se um conector chamado de BNC Barrel, cuja função era enlaçar duas extremidades machos. Por último, haviam os terminadores, cuja função era de aterramento no barramento, evitando a entrada de sinais externos no barramento.

Outra forma de conexão era adotando cabos conhecidos por AUI, estes cabos especiais funcionavam como conversores (transceivers) entre uma placa de interface fêmea, com outra placa de interface macho.

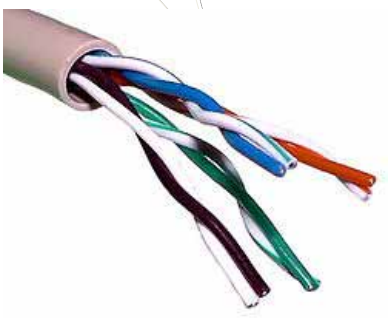




A ilustração acima da direita possibilita uma macro visão de uma rede 10Base2 (Thinnet) adotando cabeamento coaxial.

7.8.2 Par Trançado

Nos anos 90 era muito comum encontrar rede de computadores usando cabo coaxial de 50 Ohms. Isso se dava pelo fato de ser uma rede mais fácil de ser instalada pois o cabo era parecido com o cabo de antena de televisão e poderia ser instalado em qualquer local sem problemas com interferências. Com o avanço das redes de computadores, aumentando sua taxa de transferência, o cabo coaxial começou a ser substituído pelo cabo par trançado. As principais vantagens de uso do cabo par trançado são: uma maior taxa de transferência de arquivos, baixo custo do cabo e baixo custo de manutenção de rede.



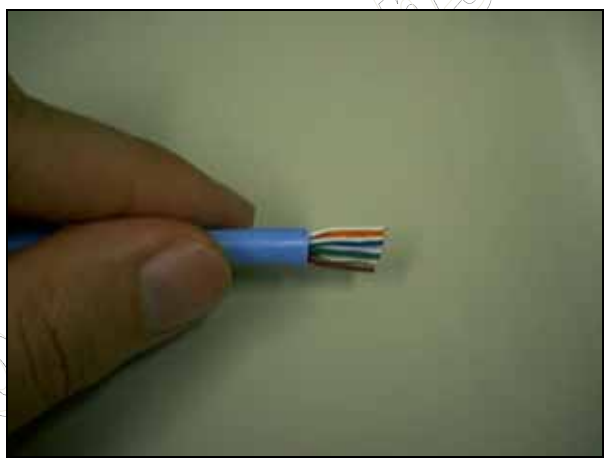
O cabeamento por par trançado (Twisted pair), ou 10BaseT, é um tipo de cabo que tem um feixe de dois fios no qual eles são entrançados um ao redor do outro para cancelar as interferências eletromagnéticas de fontes externas e interferências mútuas (linha cruzada ou, em inglês, crosstalk) entre cabos vizinhos. A taxa de giro (normalmente definida em termos de giros por metro) é parte da especificação de certo tipo de cabo. Quanto maior o número de giros, mais o ruído é cancelado. Foi um sistema originalmente produzido para transmissão telefônica analógica que utilizou o sistema de transmissão por par de fios aproveita-se esta tecnologia que já é tradicional por causa do seu tempo de uso e do grande número de linhas instaladas.

A qualidade da linha de transmissão que utiliza o par de fios depende, basicamente, da qualidade dos condutores empregados, bitola dos fios (quanto maior a bitola, menor a resistência ôhmica por quilômetro), técnicas usadas para a transmissão dos dados através da linha e proteção dos componentes da linha para evitar a indução nos condutores.

A indução ocorre devido a alguma interferência elétrica externa ocasionada por centelhamentos, harmônicos, osciladores, motores ou geradores elétricos, mau contato ou contato acidental com outras linhas de transmissão que não estejam isoladas corretamente ou até mesmo tempestades elétricas ou proximidades com linhas de alta tensão.

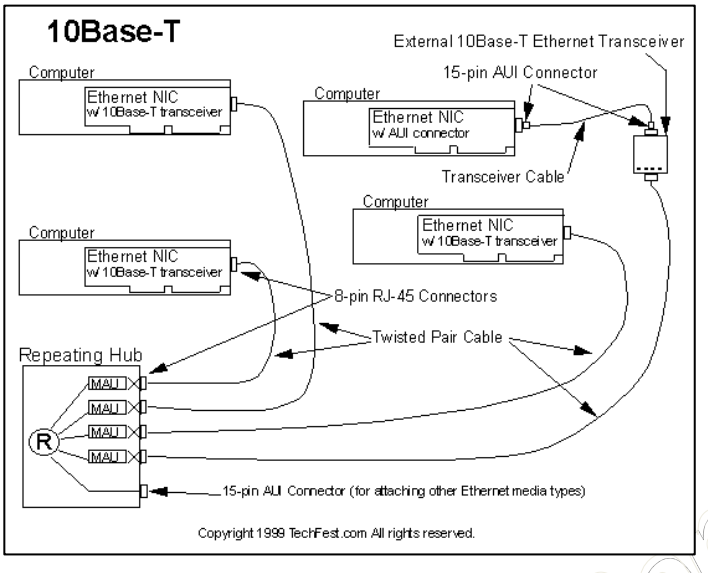
Os dois tipos mais utilizados no mercado são:

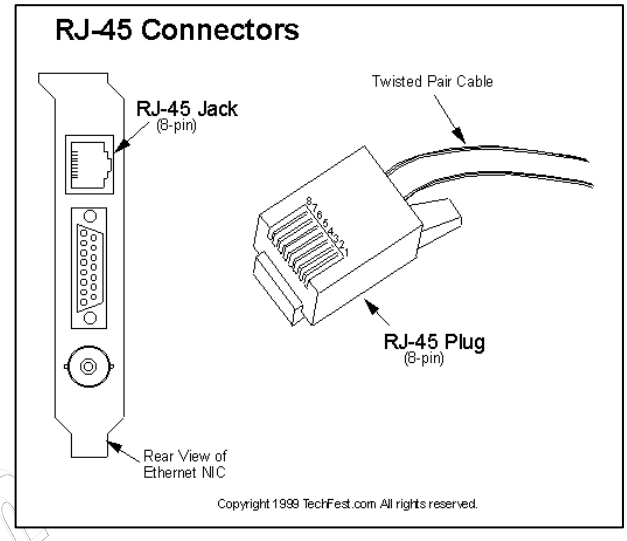
- **Unshielded Twisted Pair - UTP** ou **Par Trançado sem Blindagem**: é o mais usado atualmente tanto em redes domésticas quanto em grandes redes industriais devido ao fácil manuseio, instalação, permitindo taxas de transmissão de até 1000 Mbps com a utilização do cabo CAT 5e; é o mais barato para distâncias de até 100 metros; Para distâncias maiores emprega-se cabos de fibra óptica. Sua estrutura é de quatro pares de fios entrelaçados e revestidos por uma capa de PVC. Pela falta de blindagem este tipo de cabo não é recomendado ser instalado próximo a equipamentos que possam gerar campos magnéticos (fios de rede elétrica, motores, inversores de frequencia) e também não podem ficar em ambientes com umidade.



- **Shield Twisted Pair - STP** ou **Par Trançado Blindado** (cabo com blindagem): É semelhante ao UTP. A diferença é que possui uma blindagem feita com malha metálica, apresenta maior custo do que o UTP e o Thinnet em função dessa matéria prima, porém ainda são financeiramente mais econômicos do que a fibra ótica. Suportam uma capacidade de 260 nós contra os 1.024 do UTP, e por serem mais rígidos apresentam uma maior dificuldade de instalação. É recomendado para ambientes com relativa interferência eletromagnética, e geralmente encontrados em redes AppleTalk e TokenRing. Caso o ambiente possua umidade, grande interferência eletromagnética, distâncias acima de 100 metros ou seja exposto ao sol ainda é aconselhavel o uso de cabos de fibra ótica.

Uma rede de computadores baseada no padrão par trançado geralmente se apresentam em uma topologia física estrela, adotando como elo central o hub ou switch, e pode conviver com antigos equipamentos das redes coaxiais, adotando um conversor (transceiver) adequado.





São características desse tipo de rede: o conector RJ-45 Macho, e a interface de rede com o acesso RJ-45 Fêmea.

O cabeamento par trançado UTP foi homologado para uso nas redes locais de computadores, seus padrões, ou categorias conforme a EIA/TIA-568-A, são:

CATEGORIAS 1 e 2	Eram recomendadas para comunicação de voz e dados até **9,6Kbps**.
CATEGORIA 3	Características de desempenho para cabeamento e conexões em transmissões de dados e voz, velocidade de até **10Mbps**;
CATEGORIA 4	Características de desempenho para cabeamento e conexões em transmissões de dados e voz na velocidade de até **16Mbps**;
CATEGORIA 5	Características de desempenho para cabeamento e conexões em transmissões de dados e voz na velocidade de até **100Mbps**;
CATEGORIA 5e	(Enhanced - Melhorada), é uma melhoria das características dos materiais utilizados na categoria 5, que permite um melhor desempenho, sendo especificada até 100Mhz e na velocidade de até **1Gbps**;
CATEGORIA 6	Características para desempenho especificadas até 250Mhz e velocidades de **1Gbps até 10Gbps**.
CATEGORIA 7	Em fase de homologação.

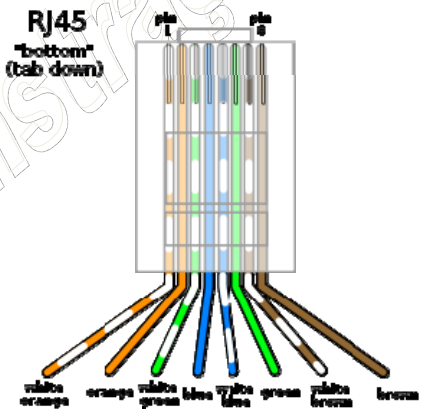
Atualmente o padrão mais utilizado é a Categoria 5e, em função de seu custo e vantagens, porém uma procura crescente pela Categoria 6 começa a surgir nos fornecedores brasileiros. Agora que já conhecemos os tipos de cabeamento par trançado, conhecemos seus padrões, vamos compreender sua montagem, item fundamental para todos os profissionais de rede.

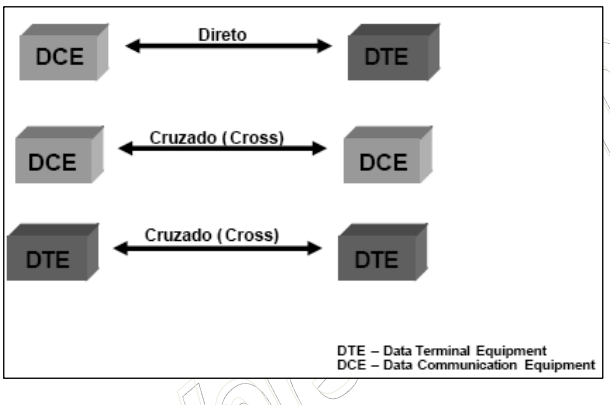
Inicialmente a IEE estipulou dois padrões para a crimpagem (confecção) do cabo ao conector RJ-45, o chamado padrão EIA/TIA-568-A (ou normal), e o padrão EIA/TIA-568-B (invertido ou cross-over):

Pino	Par 568A	Par 568B	Fio	Cor 568A Color	Cor 568B
1	3	2	tip	branco/verde	branco/laranja
2	3	2	ring	verde	laranja
3	2	3	tip	branco/laranja	branco/verde
4	1	1	ring	azul	azul
5	1	1	tip	branco/azul	branco/azul
6	2	3	ring	laranja	verde
7	4	4	tip	branco/marrom	branco/marrom
8	4	4	ring	marrom	marrom

O padrão A é certamente o mais utilizado, representando os patch-cords (cabeamento entre a estação de trabalho e a tomada de rede), patch-line (cabeamento entre o hub/switch e o patch-panel), entre outros, que veremos com maiores detalhes no tópico de cabeamento estruturado.

O padrão B, por sua vez, dá origem ao chamado cabo “crossover”. Um cabo crossover, é um cabo de rede par trançado que permite a ligação ponto-a-ponto de dois ativos de rede, como: 2 (dois) computadores pelas respectivas placas de rede sem a necessidade de um concentrador (Hub ou Switch), ou a ligação de modems entre si, ou a ligação de um computador a um roteador, ou a ligação entre concentradores.





Outra forma de compreender a funcionalidade do crossover é através da imagem ao lado, onde percebemos que o cabo crossover tem como função interligar equipamentos de mesma função, ou seja: um DCE com outro DCE ou um DTE com outro DTE.

A ligação é feita com um cabo de par trançado onde tem-se: em uma ponta o padrão T568A, e, em outra, o padrão T568B. Ou seja, inverte-se apenas uma das extremidades do cabo para se conseguir o modo crossover. Uma dica, para aqueles que não gostam de memorizar padrões, é a regra 13-26.

13 é um número fácil de recordar, para os mais supersticiosos corresponde ao número do azar. 26 é o dobro de 13. A regra 13-26 corresponde a inversão do filamento 1 com o 3 (cruz menor), e depois o filamento 2 com o 6 (cruz maior). Pronto, agora você já tem um cabo crossover. Verifique essa dica na ilustração acima.

A EIA/TIA também regula a questão do comprimento dos cabos, a fim de que a qualidade técnica seja universal. É empírico dizer que um cabo padronizado em 100 metros possa realizar 200 metros, pois vemos isso de fato nas ruas. Porém ninguém pode cientificamente afirmar que este mesmo cabo de 200 metros irá funcionar corretamente em todas as empresa e localidades do mundo, enquanto que este cabo dentro da norma de 100 metros é garantido internacionalmente para funcionamento. Dessa forma, temos uma tabela geral de comparação entre o cabeamento coaxial e o par trançado em termos de seus comprimentos.

Ethernet	Comp. Máx, do Cabo	Comp. c/ Repetidor	Topologia	Cabo	Conector	Nome Alternativo
10Base2	200m	925m	Barramento	RG-58	BNC, BNC-T, Barrel	Coaxial Fino (Thinnet)
10Base5	500m	2,5Km	Barramento	RG-11	AUX ou DIX	Coaxial Grosso (Thicknet / Ethernet)
10BaseT	100m		Física: estrela Lógica: barram.	UTP (cat-3-5)	RJ-45	Par Trançado Não Blindado

7.8.3 Fibra Ótica

A transmissão da luz pela fibra segue um princípio único, independentemente do material usado ou da aplicação: é lançado um feixe de luz numa extremidade da fibra e, pelas características ópticas do meio (fibra), esse feixe percorre a fibra por meio de reflexões sucessivas.

A fibra possui no mínimo duas camadas: o núcleo e o revestimento. No núcleo, ocorre a transmissão da luz propriamente dita. A transmissão da luz dentro da fibra é possível graças a uma diferença de índice de refracção entre o revestimento e o núcleo, sendo que o núcleo possui sempre um índice de refração mais elevado, característica que aliada ao ângulo de incidência do feixe de luz, possibilita o fenômeno da reflexão total.

As fibras ópticas são utilizadas como meio de transmissão de ondas eletromagnéticas (como a luz) uma vez que são transparentes e podem ser agrupadas em cabos. Estas fibras são feitas de plástico ou de vidro. O vidro é mais utilizado porque absorve menos as ondas eletromagnéticas. As ondas eletromagnéticas mais utilizadas são as correspondentes à gama da luz infravermelha.

O meio de transmissão por fibra óptica é chamado de "guiado", porque as ondas eletromagnéticas são "guiadas" na fibra, embora o meio transmita ondas omnidirecionais, contrariamente à transmissão "sem-fio", cujo meio é chamado de "não-guiado". Mesmo confinada a um meio físico, a luz transmitida pela fibra óptica proporciona o alcance de taxas de transmissão (velocidades) elevadíssimas, da ordem de dez elevado à nona potência a dez elevado à décima potência, de bits por segundo (cerca de 1Gbps), com baixa taxa de atenuação por quilômetro. Mas a velocidade de transmissão total possível ainda não foi alcançada pelas tecnologias existentes. Como a luz se propaga no interior de um meio físico, sofrendo ainda o fenômeno de reflexão, ela não consegue alcançar a velocidade de propagação no vácuo, que é de 300.000 km/segundo, sendo esta velocidade diminuída consideravelmente.

Cabos fibra óptica atravessam oceanos. Usar cabos para conectar dois continentes separados pelo oceano é um projecto monumental. É preciso instalar um cabo com milhares de quilômetros de extensão sob o mar, atravessando fossas e montanhas submarinas. Nos anos 80, tornou-se disponível, o primeiro cabo fibra óptica intercontinental desse tipo, instalado em 1988, e tinha capacidade para 40.000 conversas telefônicas simultâneas, usando tecnologia digital. Desde então, a capacidade dos cabos aumentou. Alguns cabos que atravessam o oceano Atlântico têm capacidade para 200 milhões de circuitos telefônicos.

Para transmitir dados pela fibra óptica, é necessário um equipamento especial chamado infoduto, que contém um componente fotoemissor, que pode ser um diodo emissor de luz (**LED**) ou um diodo ***laser***. O fotoemissor converte sinais elétricos em pulsos de luz que representam os valores digitais binários (0 e 1).

CARACTERÍSTICA	LASER	LED
Potência	Alta	Baixo
Utilização	Complexa	Simples
Velocidade	Rápida	Lento
Modo	Multimodo/Monomodo	Multimodo
Distância	Longa	Pequena
Custo	Alto	Baixo

Em Virtude das suas características, as fibras ópticas apresentam bastantes vantagens sobre os sistemas elétricos:

- Dimensões reduzidas;
- Capacidade para transportar grandes quantidades de informação (Dezenas de milhares de conversações num par de Fibra);
- Atenuação muito baixa, que permite grandes espaçamentos entre repetidores, com distância entre repetidores superiores a algumas centenas de quilômetros;
- Imunidade às interferências eletromagnéticas;
- Matéria-prima abundante;

Como desvantens, podemos citar:

- o custo, ainda muito elevando em termos de equipamentos óticos e manutenção;
- Fragilidade, acentuada pela dificuldade de fusão de fibras rompidas;
- Baixa usabilidade, requerendo que o operador cuide da dobratura máxima de 90º;

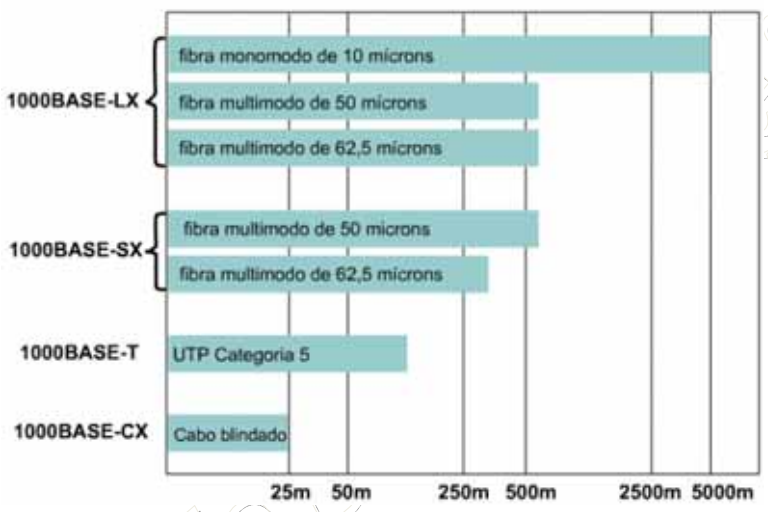
Uma característica importante que torna a fibra óptica indispensável em muitas aplicações é o fato de não ser susceptível à interferência electromagnética, pela razão de que não transmite pulsos elétricos, como ocorre com outros meios de transmissão que empregam os fios metálicos, como o cobre. Podemos encontrar aplicações do uso de fibra óptica na medicina (endoscopias por exemplo) como também em telecomunicações em substituição aos fios de cobre.

As fibras ópticas podem ser basicamente de dois modos:

- **Monomodo**:
  - Permite o uso de apenas um sinal de luz pela fibra.
  - Dimensões menores que as fibras ID.
  - Maior banda passante, de 10GHz, por ter menor dispersão.
  - Geralmente é usado laser como fonte de geração de sinal.
- **Multimodo**:
  - Permite o uso de fontes luminosas de baixa ocorrência tais como LEDs (mais baratas).

- Diâmetros grandes facilitam o acoplamento de fontes luminosas e requerem pouca precisão nos conectores.
- Muito usado para curtas distâncias pelo preço e facilidade de implementação.
- Taxa de transmissão até 1Gbps (10Gbps em fibras especiais).

Conforme o diâmetro do seu núcleo, esses modos podem alcançar maiores ou menores distância, em função das perdas por atenuações.



Por último, as redes metropolitanas ou geograficamente distribuídas, baseadas em fibra ótica, adotam uma topologia de duplo-anel, onde o segundo anel é o backup do primeiro. Essa topologia é conhecida como FDDI (*Fiber Distributed Data Interface*), um padrão definido pela ANSI em 1987. As redes FDDI adotam uma tecnologia de transmissão idêntica às das redes Token Ring, mas utilizando cabos de fibra óptica, o que lhes concede capacidades de transmissão muito elevadas (em escala até de Gigabits por segundo) e a oportunidade de se alargarem a distâncias de até 200 Km, conectando até 1000 estações de trabalho. Estas particularidades tornam esse padrão bastante indicado para a interligação de redes através de um *backbone* – nesse caso, o *backbone* deste tipo de redes é justamente o cabo de fibra óptica duplo, com configuração em anel FDDI, ao qual se ligam as sub-redes.

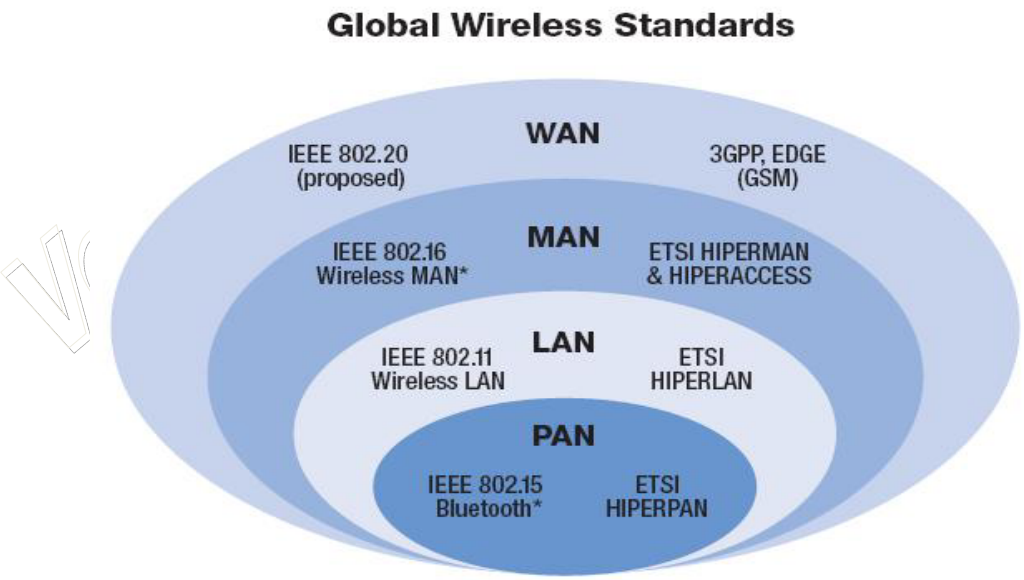
Também é importante destacar que apesar das fibras óticas serem uma alternativa as tradicionais redes par trançado e coaxial, entretanto vivemos na realidade de que as fibra óticas são utilizadas apenas em redes MAN e WAN, em função da limitação técnica de velocidade (1Gbps) e o custo. Dessa forma é comum vermos classificações topológicas de fibra como sendo exclusivamente em Anel, por causa das FDDIs, porém devemos ficar alertas para a evolução da indústria, quando os equipamentos óticos (interfaces, concentradores, conectadores e fusores) se aproximarem em custos as atuais redes de par trançado, é possível que as fibras passem a ser utilizadas nas LAN, proporcionando topologias em Barra e Estrela.

Meio de Transmissão	*Barra*	*Árvore*	*Anel*	*Estrela*
Par Trançado	X		X	X
Coaxial 50 Ohms	X		X	
Coaxial 75 Ohms	X	X		
Fibra Ótica			X	

7.8.4 Radiodifusão: redes sem fio

Como falamos no início desta seção os meios físicos se classificam em cabeados e não cabeados (sem fio), das redes cabeadas estudamos os principais padrões do mercado, como o par trançado, o caxial e a fibra ótica. Da mesma forma que as redes cabeadas apresentam tecnologias distintas, nas redes sem fio encontraremos uma gama provavelmente maior de opções. Entretanto, o mercado caminha para um padrão de nomenclatura que merece especial atenção. Quando nos referimos a redes sem fio em ambientes LAN, geralmente associamos aos padrões indoor do Wi-Fi, WiMax ou MiMo, que são concorrentes entre si para o que definimos como WLAN – Wireless LAN. E quando nos referimos a redes sem fio em ambientes WAN, geralmente associamos aos padrões outdoor dos também Wi-Fi, Wi-Max ou MiMo.

Dessa forma, as demais tecnologias como *bluetooth*, infravermelho, satélite e celular ficam designadas ao plano das telecomunicações, e não das redes de computadores. Por essa razão, optamos nessa seção em tratar com detalhes das tecnologias definidas como WLAN e, sucintamente, relacionar as demais tecnologias. Outra forma de visualizar essa justificativa, é através das distinções entre os tipos de tecnologias sem fio conforme sua abrangência geográfica.



A IEEE, *Institute of Electrical and Electronics Engineers Inc*. – www.ieee.org, é uma entidade internacional de recomendações e homologações de padrões técnicos para as tecnologias elétricas e eletrônicas. Segundo está renomada instituição, a abrangência das redes sem fios podem ser classificadas em áreas de redes pessoais (PAN), cuja abrangência não ultrapassa os 10 metros; áreas de redes locais (LAN), cuja abrangência pode variar entre 10 metros (para redes internas) até 100 metros (para redes externas); áreas de rede metropolitana (MAN), cuja abrangência varia entre 100 metros até 10 kilômetros; e áreas de redes de longo alcance (WAN), cuja abrangência varia entre 1 kilômetro até dezenas de kilômetros.

Cada classificação trás consigo representantes distintos, por exemplo, nas redes PAN são destaques as tecnologias como *bluethooth*, infravermelho e walktalkies, geralmente encontrados nos celulares e PDA para troca de arquivos. Também podemos destacar os teclados e mouses sem fio, os *home theathers* sem fio, e outros equipamentos sem fio cujas distâncias são muito próximas. Nas redes LAN destacamos as tecnologias Wi-Fi (Access points e roteadores *wirelesse* para casas e escritórios), como também o MiMo (*Multiple Inputs and Multiple Outputs*), uma variação do Wi-Fi que possibilita uma maior quantidade de dispositivos de entradas e saídas ao mesmo tempo, sem interferências entre si. Já nas redes MAN o destaque é o mesmo das redes LAN, o Wi-Fi, diferenciado apenas pelo emprego de antenas e cabeamento específicos para antenas. Em alguns casos, o mesmo equipamento adotado para uma LAN pode ser empregado para a MAN, substituindo-se apenas a antena padrão por outra de maior sensibilidade para transmissão e recepção de ondas eletromagnéticas. As WAN são marcadas basicamente pelas presenças dos celulares e dos rádios digitais. Os rádios digitais são equipamentos mais sofisticados que os Wi-Fi, possuem antenas de melhor e mais alcance também, e seus projetos geralmente só podem ser realizados por engenheiros devidamente credenciados, ao contrário das Wi-Fi cujo qualquer profissionais pode realizar o projeto, implantação e manutenção.

Os avanços nas comunicações nos últimos anos possibilitaram o surgimento de várias tecnologias, que desde então procuram atender a real necessidade de seus usuários, com a melhor qualidade possível. Nos últimos anos a comunicação sem fio ganhou um espaço considerável nas tecnologias de transmissão de dados, deixando de existir apenas nas comunicações de longa distância (feitas através de satélite), para fazer parte de ambientes locais. Essa tendência foi fortalecida pelo investimento de instituições e empresas no sentido de aplicar a transmissão sem fio em redes de computadores.

Também apostando nessa nova tecnologia, o IEEE constituiu um grupo de pesquisa para criar padrões abertos que pudessem tornar a tecnologia sem fio cada vez mais realidade. Esse projeto, denominado de Padrão **IEEE 802.11**, nasceu em 1990, mas ficou inerte por aproximadamente sete anos devido a fatores que não permitiam que a tecnologia sem fio saísse do papel. Um dos principais fatores era a baixa taxa de transferência de dados que inicialmente a tecnologia oferecia, que era em torno de Kbps.

De acordo com a elevação dessa taxa de transferência de dados que passou a atingir Mbps, a rede sem fio começou a ser vista como uma tecnologia promissora e a receber ainda mais investimentos para a construção de equipamentos que possibilitassem a comunicação sem fio entre computadores.

Atualmente o foco das redes de computadores sem fio (Wireless) se encontra no contexto das redes locais de computadores **(Wireless Local Area Network - WLAN)**, tanto em soluções proprietárias como no padrão do IEEE. Primeiramente foram colocados em prática alguns padrões proprietários, através de empresas como *IBM, CISCO, Telecom e 3COM*. Hoje essas e outras empresas baseiam seus produtos no padrão do IEEE, devido às inúmeras e já conhecidas vantagens que o padrão aberto oferece: interoperabilidade, baixo custo, demanda de mercado, confiabilidade de projeto, entre outras.

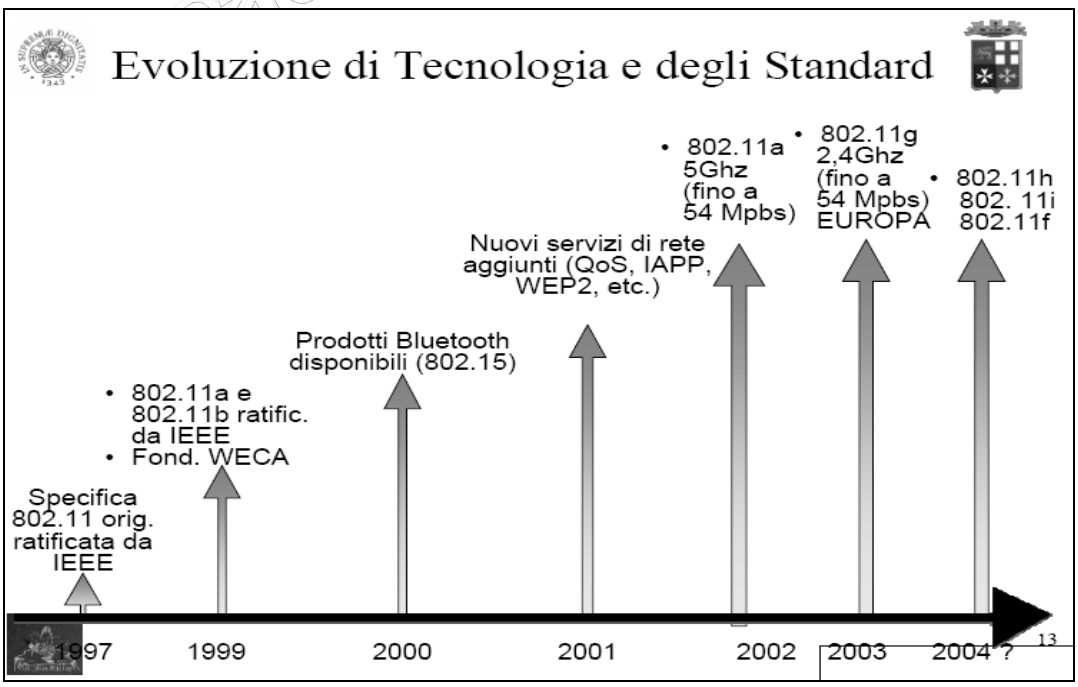
Fora das redes de computadores, muitas tecnologias sem fio proprietárias têm sido usadas para possibilitar a comunicação entre dispositivos sem fio. Essas tecnologias têm o propósito de permitir o controle remoto de equipamentos domésticos e interligar os periféricos (teclado, mouse, impressoras, etc) aos computadores, eliminando os fios e tornando mais flexível e prático o uso desses equipamentos.

O padrão IEEE 802.11 define basicamente uma arquitetura para as WLANs que abrange os níveis físicos e de enlace. No nível físico são tratadas apenas as transmissões com freqüência de rádio *(RF), que operam em 2.4GHz,* e infravermelho *(IR)*, embora outras formas de transmissão sem fio possam ser usadas, como microondas e laser, por exemplo.

O padrão IEEE 802.11 possibilita a transmissão de dados numa velocidade de 1 (obrigatório) à 2Mbps (opcional), e especifica uma arquitetura comum, métodos de transmissão, e outros aspectos de transferência de dados sem fio, permitindo a interoperabilidade entre os diversos produtos WLAN.

Apesar da significante elevação da taxa de transferência de dados que subiu de algumas poucas dezenas de kilobits por segundo para 2Mbps, as WLANs não atendiam satisfatoriamente a necessidade de banda das empresas. Com isso, o IEEE investiu no melhoramento do padrão 802.11 (que passou a ser chamado de *802.11b, ou Wi-Fi*), com a mesma arquitetura e tecnologia, mas com taxa de transferência de dados maiores, entre **5** e **11 Mbps**, automaticamente selecionados conforme a distância entre os equipamentos, e impulsionando de vez a tecnologia e estimulando as comunidades científica e industrial a padronizarem, projetarem e produzirem produtos para essas redes.

A IEEE também investiu em um melhoramento através de um novo espectro de freqüência, o 5GHz. O IEEE 802.11a, ou também conhecido como WiMax (redefinido como 802.16). Esta tecnologia consegue alcançar taxas de transmissão de dados na ordem de quatro vezes o 802.11b, ou seja 27Mbps, além de suportar uma maior quantidade de equipamentos simultâneos sem interferências. Entretanto, esta tecnologia não apresenta compatibilidade com as anteriores, fazendo com que a IEEE voltasse a investir na freqüência de 2.4GHz. O IEEE 802.11g é o aprimoramento do 802.11b realizando taxas de transferências maiores, de 54Mbps, e sendo totalmente compatível com os padrões 802.11 e 802.11b, porém não compatível com o 802.11a.



Como pode ser observado no gráfico acima, após a publicação do padrão 802.11b surgem respectivamente: 802.15, 802.11a (54 Mbps), 802.11g, entre outros. Um detalhe especial é que cada um desses padrões tratam de um determinado assunto em específico, por exemplo, os padrões 802.11a,

802.11b e 802.11g descrevem os padrões de velocidades entre dos espectros de freqüência da tecnologia. Os padrões 802.11h, 802.11i e 802.11f descrevem aspectos de interoperabilidade, segurança e qualidade de serviço em redes sem fio, respectivamente. Entre outros padrões que existem e que descrevem maiores detalhes sobre o uso das redes sem fio.

Na próxima tabela apresentamos um breve resumo sobre a evolução e característica dos principais padrões de velocidades para as redes sem fio.

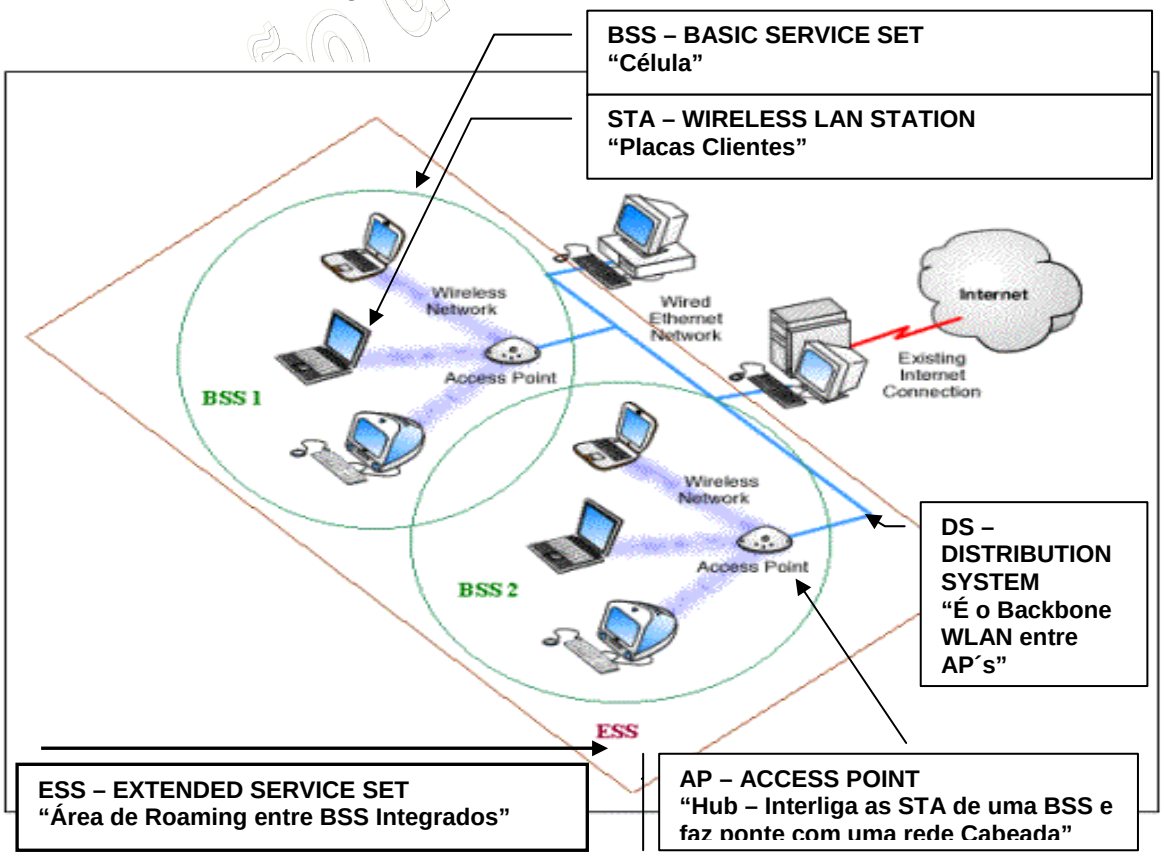
	802.11	**802.11a**	**802.11b (+)**	**802.11g**
Padronizado em	Julho 1997	Setembro 1999	Setembro 1999	Junho 2003
Freqüências de operação (Protocolo)	2.4-2.4835 GHz (DSSS e FHSS)	5.15-5.35 GHz (OFDM-UNII) 5.725-5.825Ghz (OFDM-ISM)	2.4-2.4835GHz (DSSS)	2.4-2.4835GHz (DSSS e OFDM)
Taxa de Transmissão por canal (Mbps)	2, 1	54, 48, 36, 24, 18, 12, 9, 6	22,11, 5.5, 2, 1	54, 36, 33, 24, 22, 12, 11, 9, 6, 5.5, 2,1
Tipos de Modulação (Taxa de Transmissão em Mbps / Protocolos de Modulação)	DQPSK (2/DSSS) DBPSK (1/DSSS) 4GFSK (2/FHSS) 2GFSK (1/FHSS)	BPSK (6,9) QPSK (12,18) 16QAM (24,36) 64QAM (48,54)	PBCC(22,11,5.5,2,1) DQPSK/CCK (11,5.5) DQPSK (2) DBPSK (1)	OFDM/CCK (6,9, 12,18,24,36,48,54) OFDM (6,9,12,18, 24,36,48,54) DQPSK/CCK (22, 33, 11, 5.5) DQPSK (2 Mbps) DBPSK (1 Mbps)
Compatibilidade	802.11	Wi-Fi5/WiMax	Wi-Fi	Wi-Fi a 11Mpbs e menores

As redes sem fio apresentam protocolos de comunicação diferente para suportaram abrangências e velocidades diferentes. Ou seja, quanto mais próximo estiverem os equipamentos uns dos outros, maior será a velocidade de acesso, porém, se os equipamentos estiverem muito distantes uns dos outros, então o equipamento automaticamente identifica um protocolo de longa distância, porém cuja velocidade serja reduzida.

As redes wireless possuem capacidade de substituição ou de ser adicionada aos sistemas com fio já existentes (LANs), passando a ser uma solução bastante interessante para as organizações, pois desta forma os pontos que necessitam de mobilidade são conectados à rede pelo meio ”Wireless”e as estações fixas são ligadas à rede via cabo.

Para que se entenda melhor uma arquitetura wireless é necessário que alguns conceitos sejam descritos:

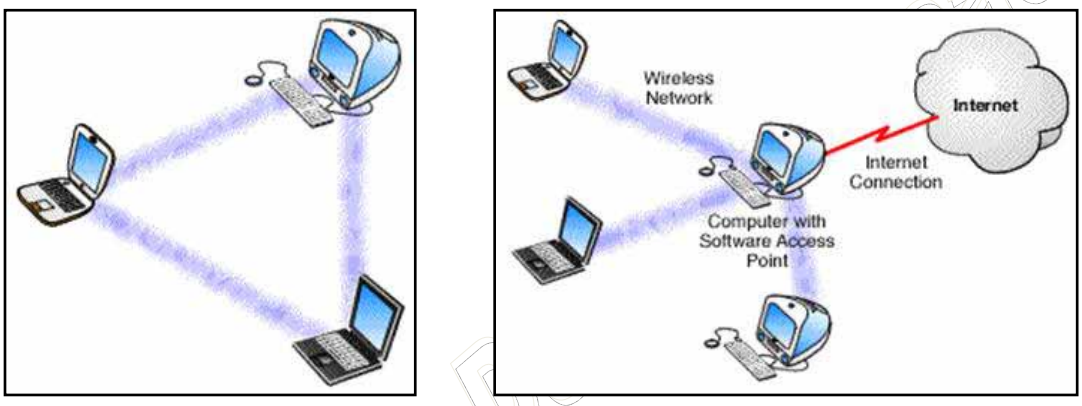
- **BSS - Basic Service Set** - corresponde a uma célula de comunicação wireless. É o raio de comunicação de uma rede wireless, ou seja, a região de transmissão;
- **STA - Stations** - são as estações de trabalho que se comunica entre si dentro da BSS. Podem ser dispositivos portáteis ou desktops, sendo que para o primeiro será necessário o uso de um cartão PC CARD, que geralmente já estão inclusos na sua arquitetura através de seus fabricantes, no segundo será necessário a aquisição de uma placa com conector PC CARD, além do cartão.
- **AP - Access Point** - funciona como uma bridge ou hub entre a rede wireless e a rede tradicional, geralmente o padrão ethernet. Coordena a comunicação entre as STA dentro da BSS . Poder ser uma antena de comunicação.
- **ESS - Extended Service Set** - consiste de várias células BSS vizinhas que se interceptam e cujos AP estão conectados a uma mesma rede tradicional. Nestas condições uma STA pode movimentar-se de um BSS para outro permanecendo conectada à rede. Este processo é denominado **Roaming**.



As redes WLAN operam de dois modos distintos:

- **Ad Hoc** - é um sistema onde as comunicações são estabelecidas entre várias estações de uma mesma área (célula), sem o uso de um ponto de acesso ou servidor e sem a necessidade de infra-estrutura (rede ponto-a-ponto).
- **Infra-estrutura** - onde possui a presença de um AP (Cliente/Servidor).

A figura a seguir apresenta um exemplo de uma rede local sem fio Ad Hoc e uma rede com infra-estrutura, respectivamente:



Numa Rede sem Fio os dados são enviados em canais de *freqüência de rádio, infravermelho ou laser*.

O infravermelho é pouco usado. Sua faixa de freqüência fica logo abaixo da freqüência da luz visível. Os sinais transmitidos devem ser bem fortes, de alta intensidade para não permitir a interferência da luz externa. Pode-se conseguir altas taxas de transmissão chegando em 10 Mbps. A distância máxima de comunicação não ultrapassa uns 30 metros mesmo com dispositivos bem potentes da atualidade.

Pode-se utilizar a transmissão por infravermelho com feixe direto (linha de visada desobstruída), semelhante à comunicação dos controles remotos das televisões caseiras, ou com radiação a todas as direções por reflexão em superfícies e teleponto óptico (lentes) de banda larga.

Já o laser pode alcançar distancias de 200 a 300 metros com visada direta. Ele pode ser utilizado para conectar duas Redes sem Fio, cada uma cobrindo, por exemplo, um prédio.

As freqüências de rádio (radiodifusão) são as mais utilizadas em redes de computadores. Por sua natureza, ela é adequada tanto para legações ponto a ponto quanto para ligações multipontos. As Redes sem Fio, baseadas em radiodifusão, são uma alternativa viável onde é difícil, ou mesmo impossível, instalar cabos metálicos ou de fibra óptica. Seu emprego é particularmente importante para comunicações entre computadores portáteis em um ambiente de rede local móvel.

A radiodifusão também é utilizada em aplicações onde a confiabilidade do meio de transmissão é requisito indispensável. Um exemplo drástico seria em aplicações bélicas, onde, por exemplo, o rompimento de um cabo poderia paralisar todo um sistema de defesa.

O Padrão IEEE 802.11 trata da tecnologia sem fio enfocando as redes locais sem fio (WLAN). Essas redes basicamente utilizam radiofreqüência para a transmissão de dados, através de duas técnicas

conhecidas como **DSSS** (*Direct Sequence Spread Spectrum*) e **FHSS** (*Frequency Hopping Spread Spectrum*), codificando dados e modulando sinais de modos diferentes para equilibrar velocidade, distância e capacidade de transmissão. A escolha da técnica DSSS ou FHSS dependerá de vários fatores relacionados com a aplicação dos usuários e o ambiente onde a rede operará.

As especificações FHSS e DSSS operam na freqüência de 2,4 GHz denominada banda **ISM** (*Industrial Scientific and Medical*) cujo uso é liberado sem necessidade de licenciamento.

Para a transmissão em radiofreqüência são usadas as técnicas DSSS e FHSS. Essas técnicas transmitem os quadros de dados enviando-os por vários canais disponíveis dentro de uma freqüência, ao invés de usar um único canal, possibilitando, dessa forma, a transmissão simultânea de vários quadros.

A técnica DSSS distribui o sinal em cima de uma gama extensiva da faixa de freqüência e reorganiza os pacotes no receptor. A técnica FHSS envia segmentos curtos de dados que são transmitidos através de freqüências específicas, controlando o fluxo com o receptor, que negocia velocidades menores comparadas às velocidades oferecidas pela técnica DSSS, mas menos suscetíveis a interferências.

O padrão 802.11 usa as duas técnicas, enquanto que outras tecnologias, como o *HomeRF e Bluetooth*, usam apenas a técnica FHSS, que é mais eficiente para ambientes que possuem outros tráficos de rádio, como áreas públicas abertas, por exemplo.

As WLANs baseadas em radiofreqüência usam freqüências de 900MHz, 2.4GHz e 5GHz. Quanto maior a freqüência maior é a quantidade de informação que um dispositivo pode enviar num canal.

As primeiras WLANs operavam na freqüência de 900MHz, atingindo uma taxa de 256Kbps. O padrão IEEE 802.11 aumentou a taxa de transmissão para 1Mbps, usando a técnica FHSS, e posteriormente para 2Mbps, usando a técnica DSSS, trabalhando na freqüência de 2.4GHz.

A maioria das empresas optou pela técnica DSSS porque oferece freqüências mais altas do que a FHSS.

As dúvidas e receios dos usuários sobre a segurança em comunicação sem fio, é um dos maiores obstáculos para os vendedores de redes sem fio e tem contribuído para o lento desenvolvimento do mercado. Compradores em potencial comparam sem fio com broadcasting e temem que dados privados possam ser, e realmente são, facilmente acessados por curiosos.

As comunicações sem fio 802.11 não pode ser recebido, e muito menos decodificado,por simples rastreadores, receptores de ondas curtas, etc. Isto está baseado na concepção comum de que comunicações sem fio não podem ser acessados por qualquer um. Porém, invasão é possível usando equipamentos especiais. Para proteger contra qualquer falha de segurança em potencial, o padrão 802.11 inclui duas funções, **WEP e WPA.**

WEP (*Wired Equivalent Privacy*) é uma forma de criptografia que provê privacidade comparável ao existente em uma rede local tradicional. É baseado em proteger o dado transmitido pelo meio de radiofreqüência usando uma chave de 64 bits e o algoritmo de criptografia RC4. WEP, quando habilitado, apenas protege a informação do pacote de dados, e não protege o cabeçalho da camada física de modo que outras estações na rede possam receber os dados de controle necessários para gerenciar a rede. No entanto, as outras estações não podem descriptografar a partição de dados do pacote. O WEP surge juntamente com o padrão 802.11b, porém desde seu surgimento diversas formas de decodificar sua cifra criptográfica foram apresentadas.

WPA (Wi-Fi Protected Access) é o substituto do WEP, lançado em 2002 em conjunto com o novo padrão de criptografia mundial, o AES (Advanced Encryption Standard). Sua cifra é inviolável, garantida pelas maiores instituições internacionais de criptografia. Por ainda ser recente seu lançamento, apresenta incompatibilidades com equipamentos antigos ou que foram desenvolvidos para operar exclusivamente com WEP, entretanto sua robustez e confiabilidade já conquistaram os mercados, que cada vez mais tende a ter no WPA o protocolo padrão de segurança.

7.9 INSTALAÇÕES FÍSICAS E CABEAMENTO ESTRUTURADO

Edifícios, casas e prédios são construídos tendo em mente várias décadas de uso, senão séculos, como demonstram igrejas, pontes e palácios na Europa. O projeto arquitetônico pode ficar desatualizado e o uso da edificação pode mudar com o tempo (escolas transformadas em bibliotecas, casas transformadas em escritórios), mas a edificação em si deve resistir à passagem dos anos sem que precise ser reconstruída continuamente.

A estrutura elétrica, entretanto, nem sempre vislumbra estas várias décadas de uso, especialmente a de comunicação. Isto quando a edificação não é antiga a ponto de possuir apenas uma estrutura elétrica mínima, com basicamente pontos de iluminação. Mesmo que o material usado na estrutura da instalação elétrica preveja vários anos de vida útil, o projeto da instalação em si raramente prevê modificação de uso da edificação, adição de maior número de equipamentos elétricos, ou mesmo adição de equipamentos elétricos que ainda nem existem. O que se vê, então, comumente, são dificuldades em instalar novas linhas de telefone em uma residência ou escritório, ou um aparelho de televisão em um recinto onde não havia, um aparelho de fax, um ponto de conexão à Internet. Freqüentemente essas operações necessitam de uma reforma para ampliação do cabeamento, passando por quebra de paredes e pisos, pintura e troca de cabeamento já existente, gerando custo e transtorno.

Por outro lado, vários sistemas de comunicação diferentes exigem vários sistemas de cabeamento diferentes e dedicados. Fios telefônicos para voz, cabos coaxiais para tv, cabos multivias para dados.

O cabeamento estruturado propõe-se a resolver esses contratempos, provendo padrões para os cabos e conexões, de modo que, com a adição de equipamento adicional, possam suportar todos ou praticamente todos os diferentes tipos de sistema de comunicação hoje em uso, numa mesma instalação.

7.9.1 Definição e Características

Cabeamento estruturado é um cabeamento de baixa corrente e tensão para uso integrado em comunicações de voz, dados, controles prediais e imagem, preparado de tal maneira que atende aos mais diversos tipos e layouts de instalação, por um longo período de tempo, sem exigir modificações físicas da infra-estrutura. A idéia é que este cabeamento proporcione ao usuário uma tomada universal, onde ele possa conectar diferentes aplicações como computador, telefone, fax, rede local, TV a cabo, sensores, alarme, etc. Isto se contrapõe ao conceito de cabeamento dedicado, onde cada aplicação tem seu tipo de cabo e instalação. Assim, sinal de tv requer cabos coaxiais de 75 ohms e conectores e painéis específicos;

o sistema de telefonia requer fios apropriados, tomadas e painéis de blocos específicos; redes de computadores usam ainda cabos multivias dedicados. Isso resulta em diversos padrões proprietários ou não de cabos, topologias, conectores, padrões de ligações, etc. O conceito de cabeamento estruturado surge como resposta com o intuito de padronizar o cabeamento instalado em edifícios comerciais ou residenciais, independente das aplicações a serem usadas nos mesmos.

O cabeamento estruturado provavelmente originou-se de sistemas telefônicos comerciais, onde o usuário constantemente mudava sua posição física no interior de uma edificação. Projetou-se um cabeamento de modo a existir uma rede horizontal fixa, ligada a uma central de distribuição, onde cada ponto podia ser ativado ou desativado facilmente. Um ponto de tomada podia ser rapidamente alternado ou deslocado por meio de uma troca de ligações. O sistema evoluiu para que diversos tipos de redes pudessem ser interligados, mantendo o cabeamento horizontal e tornando as tomadas de uso múltiplo.

A solução do cabeamento estruturado prevê a instalação de um cabo e um tipo de conector padrão, e equipamentos adicionais para suporte a diferentes tipos de sistemas. Isto é conhecido como *Cabeamento Genérico*.

Para assegurar flexibilidade, é de interesse que este cabeamento genérico esteja instalado e pronto para uso em todos os locais possíveis em um determinado local ou edificação. Isso permite, por exemplo, a expansão ou mudança de um departamento em um escritório para outras dependências com o mínimo de transtorno e custo. Esta tática é conhecida em inglês como *Flood Wiring* , e que consiste no espalhamento de conexões por todo o recinto (cerca de duas conexões por cada 3 m^2 de recinto).

De modo a permitir que diferentes tomadas possam ser usadas para sistemas distintos, um painel especial conhecido como *Patch Panel* é utilizado.

Estes três atributos: *Cabeamento Genérico, Flood Wiring* e *Patch Panels* são as características essenciais de um sistema de cabeamento estruturado.

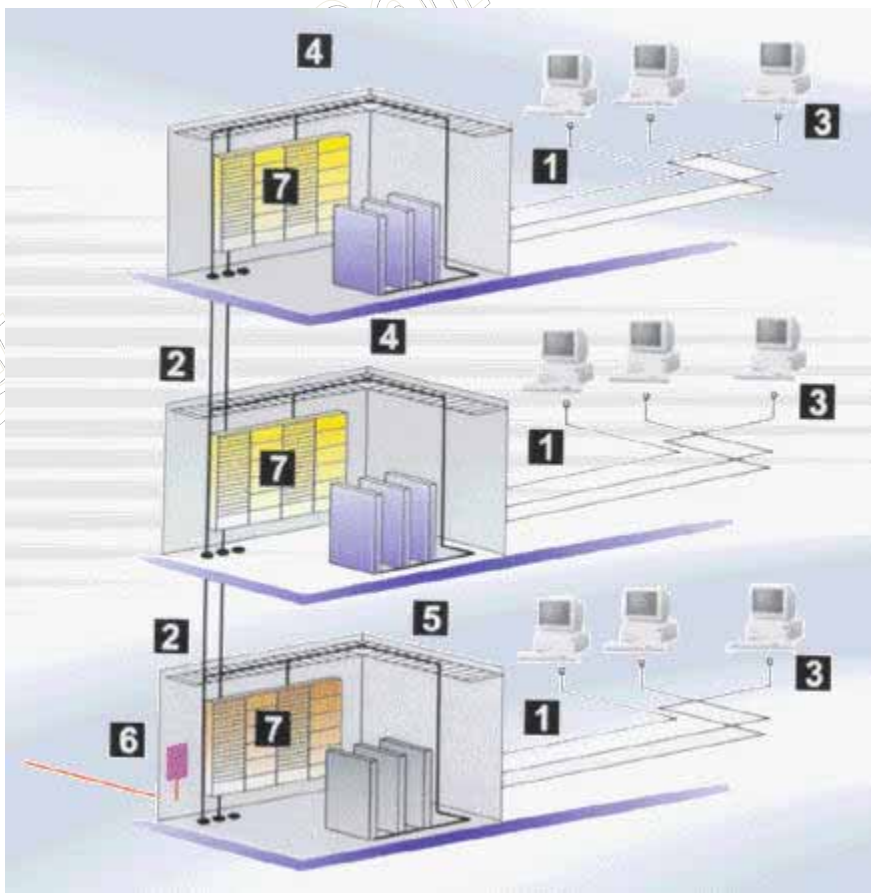
7.9.2 Estrutura e Topologia

De acordo com as normas ANSI/TIA/EIA-568-A e ANSI/TIA/EIA-606, a instalação de um cabeamento divide-se em basicamente oito elementos:

1. Cabeamento Horizontal: são os cabos que ligam o painel de distribuição até o ponto final do cabeamento (tomadas). Estes cabos formam um conjunto permanente e são denominados cabos secundários.
2. Cabeamento Vertical ou *backbone*: conjunto permanente de cabos primários, que interligam a sala de equipamentos aos TC's e pontos de Entrada (EF's).
3. Posto de Trabalho ou *work area*: ponto final do cabeamento estruturado, onde há uma tomada fixa para a conexão do equipamento. Se o local de instalação não é um escritório, ou seja, é uma edificação residencial, o "posto de trabalho" é qualquer ponto final onde há uma tomada.
4. Armários de Telecomunicações ou *Telecommunications Closets (TC's)*: espaço para acomodação dos equipamentos, terminações e manobras de cabos. Ponto de conexão entre o backbone e o cabeamento horizontal.

5. Sala de Equipamentos ou *Equipment Room (ER)*: recinto onde se localizam os equipamentos ativos do sistema bem como suas interligações com sistemas externos. Ex.: central telefônica, servidor de rede de computadores, central de alarme. Este recinto pode ser uma sala específica, um quadro ou *shaft*. Costuma-se também instalar neste local o principal painel de manobras ou *Main Cross-Connect*, que pode ser composto de *patch-panels*, blocos 110, blocos de saída RJ-45 ou distribuidores óticos.
6. Entrada da Edificação ou *Entrance Facilities (EF)*: ponto onde é realizado a interface entre o cabeamento externo e o interno da edificação para os serviços disponibilizados.
7. Painéis de Distribuição ou *Cross-Connect*: recebem, de um lado, o cabeamento primário vindo dos equipamentos, e de outro o cabeamento horizontal, que conecta as tomadas individuais. A ativação de cada tomada é feita no painel de distribuição, por intermédio dos *patch-panels*.
8. *Patch-panels*: painéis formados por conjuntos gêmeos de portas, que recebem a conexão de um cabo por um lado, conectam este cabo ao painel gêmeo por meio de um *patch-cord*, e que finalmente recebe a conexão de um outro cabo. Através da manobra com os *patch-cords*, as conexões podem ser refeitas e realocadas com velocidade e simplicidade.

Os *cross-connects* e os TC's podem ser aglutinados numa só peça.



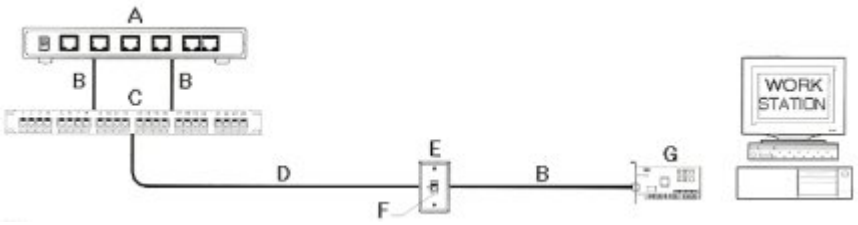


Figura : Patch-panel. (a) hub; (b) patch-cord; (c) patch-panel; (d) cabo horizontal; (e) espelho de tomada; (f) conector; (g) placa de rede.

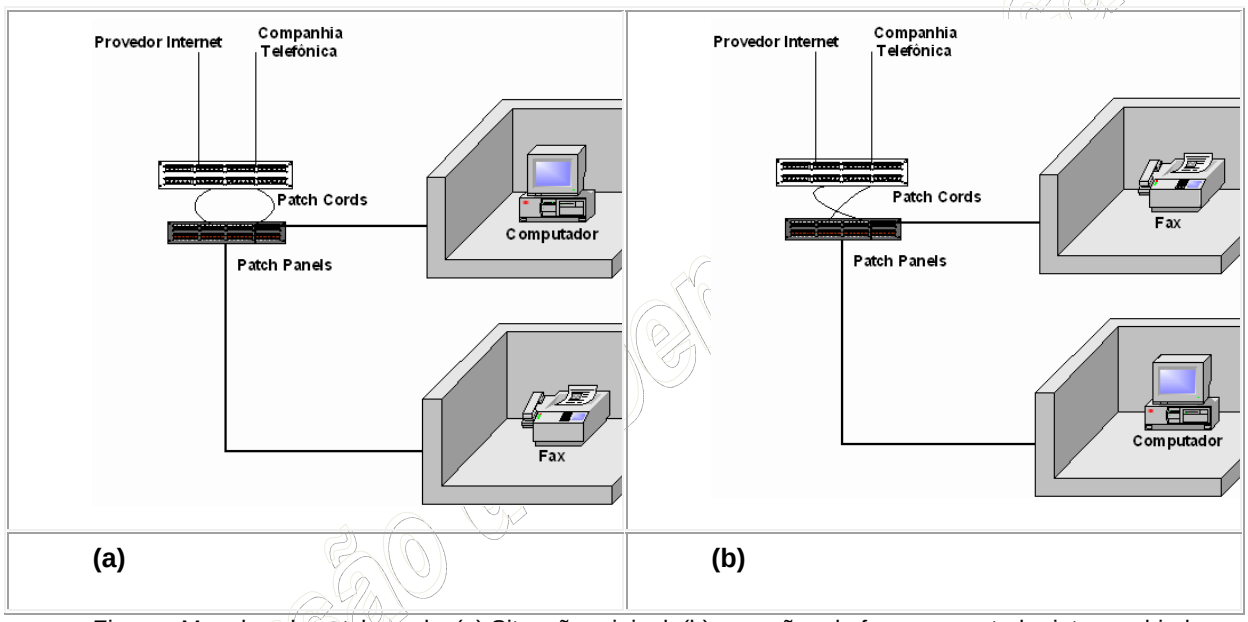


Figura : Manobra de patch-cords. (a) Situação original; (b) conexões de fax e computador intercambiadas.

A Figura acima ilustra uma manobra do *patch-panel*. Em um determinado recinto, há um computador conectado à Internet. Em outro, há um aparelho de fax. Decide-se intercambiar estes equipamentos, colocando o computador no recinto do fax e vice-versa. As conexões para cada equipamento são rapidamente e facilmente configuradas bastando trocar os *patch-cords* correspondentes de posição no *patch-panel*.

Uma instalação típica de cabeamento estruturado consiste em tomadas para o usuário com conectores do tipo RJ-45. Estas tomadas contém um ou dois conectores RJ-45, cada, montadas na parede ou ainda em caixas no piso.

Cada cabo vindo dessas tomadas para o usuário é então trazido para os Telecommunications Closets (TC's) usando cabos de quatro pares de fios trançados (cabeamento horizontal). Na maioria dos casos, usa-se cabos Categoria 5e para o cabeamento horizontal, podendo estes cabos ser UTP ou STP. Os cabos são conectados na tomada através de um dispositivo chamado IDC (Insulation Displacement Connection).

Figura : Pequeno gabinete ou TC residencial.

No sistema de cabeamento estruturado, no cabo horizontal trafegam todos os serviços, sejam voz, rede, vídeo, controle ou outras aplicações. Se os requerimentos de uso mudarem, o serviço provido para as tomadas correspondentes pode ser mudado bastando configurar os patch-cords devidos no painel. Se necessário, um adaptador é usado na tomada para converter ou compatibilizar o serviço. Por exemplo, um balun de conversão para vídeo.

A filosofia de "flood wiring" consiste em instalar tomadas no recinto de acordo com uma densidade, ou área do recinto, ao invés de focar na posição final do usuário. Isso permite maior flexibilidade, pois quando mudanças são feitas no layout, não é preciso re-cabear o recinto.

No TC, os cabos individuais de par trançado vindos das tomadas são terminados nos patch-panels, através dos IDC. Os patch-panels contém conectores RJ-45 na frente, para conexão dos patch-cords. Os patch-panels são comumente montados em racks apropriados e afixados na parede ou piso.

Na instalação de cabeamento estruturado, não se conecta diretamente um equipamento que provê um serviço ou sinal (equipamento ativo) ao usuário. Por exemplo, não se conecta diretamente um PC a um hub. Conforme prescrevem as normas de cabeamento estruturado, o equipamento ativo deve ser conectado ao painel distribuidor, e este (através dos patch-panels) ser conectado a uma tomada. Isto torna o sistema independente e aberto, configurando-lhe agilidade.

No Brasil, a norma mais conhecida para cabeamento estruturado é a ANSI/TIA/EIA 568-A, fruto do trabalho conjunto da TIA (Telecommunications Industry Association) e a EIA (Electronics Industries Association). Esta norma prevê os conceitos apresentados anteriormente e é complementada por outras normas. A Tabela 1 contém as normas observadas na instalação de cabeamento estruturado. A Tabela 2 traz as categorias dos cabos UTP e suas respectivas larguras de banda, que também são usados como diretrizes para os projetos e instalação.

Norma	**Tema**
ANSI/TIA/EIA 568-A	Padrões de Cabeamento
ANSI/TIA/EIA 569-A	Infra-estrutura
ANSI/TIA/EIA 570-A	Cabeamento Residencial
ANSI/TIA/EIA 606	Administração
ANSI/TIA/EIA 607	Aterramento

Tabela 1: Normas para cabeamento estruturado.

Categoria	Largura de Banda
1 e 2	Até 9,6Kbps
3	Até 10Mbps
4	Até 16Mbps
5	Até 100Mbps
5e	Enhanced - Até 100Mbps (menos ruídos)
6	De 1Gbps até 10Gbps

Tabela 2: Categoria dos Cabos UTP/STP.

Pode-se citar alguns benefícios proporcionados pela utilização de cabeamento estruturado, em lugar de cabeamento convencional:

- Flexibilidade: permite mudanças de layout e aplicações, sem necessidade de mudar o cabeamento.
- Facilidade de Administração: as mudanças de aplicações, manutenção e expansão são feitas por simples trocas de *patch-cords* ou instalação de poucos equipamentos adicionais.
- Vida Útil: o cabeamento tipicamente possui a maior expectativa de vida numa rede, em torno de 15 anos. O cabeamento estruturado permite a maximização dessa vida útil, utilizando-se do mesmo cabo para transportar várias tecnologias de comunicação ao mesmo tempo, e também prevê a implementação de tecnologias futuras, diferentes das utilizadas no período da instalação.
- Controle de Falhas: Falhas em determinados ramos do cabeamento não afetam o restante da instalação.
- Custo e Retorno sobre Investimento (ROI – *Return of Investment*): O Sistema de Cabeamento Estruturado consiste em cerca de 2 a 5% do investimento na confecção de uma rede. Levando em conta a vida útil do sistema, este certamente sobreviverá aos demais componentes dos serviços providos, além de requerer poucas atualizações com o passar do tempo. Ou seja, é um investimento de prazo de vida muito longo, o que o torna vantajoso.

A demanda por serviços de comunicação, tais como voz, imagem, dados e controles prediais tem saboreado um crescimento constante, ainda que, no período entre os anos de 1999-2001, a oferta tenha sido muito maior, acarretando complicações financeiras particularmente para as empresas de telecomunicação. Esta demanda é verdadeira tanto em empresas como em residências, com a instalação de mais de uma linha telefônica, ou a instalação de telefones em vários cômodos e pontos de interligação de computadores (rede de computadores) em vários cômodos e entre residências num mesmo condomínio.

Para as empresas, a comunicação é vital para a operação dos negócios, seja voz, seja dados, e principalmente dados. Os novos prédios comerciais têm freqüentemente adotado métodos de controle predial (edifícios inteligentes), como forma de aperfeiçoar e melhorar segurança e uso de eletricidade, bem como o conforto.

Assim sendo, o sistema de cabeamento estruturado surge como opção óbvia para o projeto de edificações, em lugar do cabeamento convencional, onde cada sistema ou tecnologia exige seu cabeamento próprio. O cabeamento estruturado é flexível, pois permite a agregação de várias tecnologias sobre uma mesma plataforma (ou cabo); é de fácil administração, pois qualquer mudança não passa pela troca dos cabos, e sim por configuração em painéis próprios; tem relação investimento/benefício excelente, pois prevê longa vida útil, com suporte a tecnologias futuras com pouca ou nenhuma modificação, e permite modificações de layout ou de serviços providos com a simples alteração de conexões no painel.

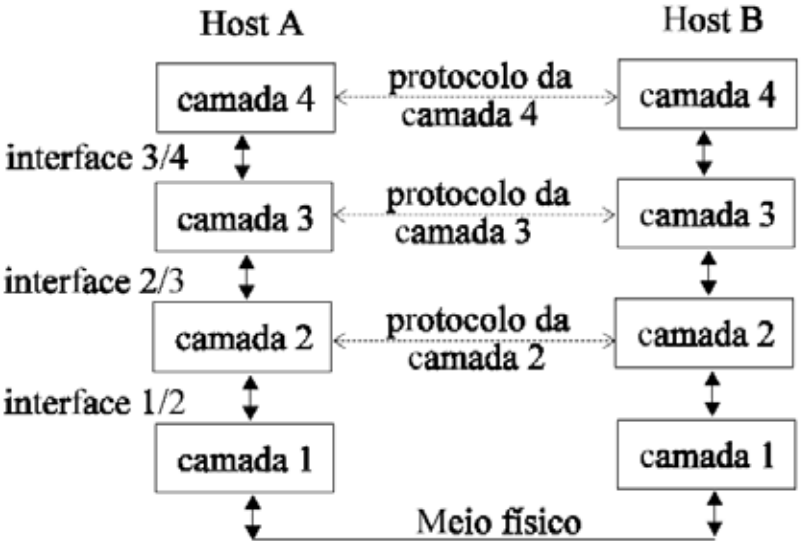
8 COMPETÊNCIA 3 – MODELOS DE REFERÊNCIA EM REDES DE COMPUTADORES

8.1 ARQUITETURAS DE REDES DE COMPUTADORES

A tarefa de permitir a comunicação entre aplicações executando em máquinas distintas envolve uma série de detalhes que devem ser cuidadosamente observados para que esta comunicação ocorra de maneira precisa, segura e livre de erros. Por exemplo, detalhes de sinalização dos bits para envio através dos meios de transmissão; detecção e correção de erros de transmissão (pois a maioria dos meios de transmissão é passível de interferências); roteamento das mensagens, desde sua origem até o seu destino, podendo passar por várias redes intermediárias; métodos de endereçamento tanto de hosts quanto de aplicações; cuidar da sintaxe e semântica da informação, de modo que quando uma aplicação transmite um dado do tipo inteiro, a aplicação destino possa entendê-lo como do tipo inteiro; etc.

Para reduzir a complexidade de projeto, a maioria das redes de computadores é estruturada em **camadas** ou **níveis**, onde cada camada desempenha uma função específica dentro do objetivo maior que é a tarefa de comunicação. As camadas são construídas umas sobre as outras e cada camada oferece seus serviços para as camadas superiores, protegendo estas dos detalhes de como os serviços oferecidos são de fato implementados.

A **camada N** em uma máquina, para desempenhar suas funções estabelece uma conversação com a **camada N** em outra máquina. As regras utilizadas nesta conversação são chamadas de **protocolo da camada N**. As funções de cada camada são executadas por **entidades** (processos, que podem ser implementados por software ou por hardware). Entidades que executam em camadas correspondentes e em máquinas distintas são chamadas de **processos pares** (*peers*). São os processos pares que se comunicam utilizando o protocolo de sua camada. A figura a seguir ilustra estes conceitos para uma rede estruturada em 4 camadas.



Na verdade, nenhum dado é transferido diretamente da camada N de uma máquina para a camada N de outra máquina. Em vez disso, cada camada passa dados e informações de controle para a camada

imediatamente abaixo, até encontrar o meio físico, através do qual a comunicação de fato ocorre. Na máquina destino a mensagem percorre o caminho inverso, da camada mais inferior para a mais superior, com cada camada retirando e analisando as informações de controle colocadas pela sua camada correspondente na máquina origem. Após esta análise a camada decide se passa o restante dos dados para a camada superior. Estas informações de controle correspondem ao protocolo da camada e também são conhecidos como ***header* do protocolo**.

Para ilustrar o conceito de comunicação através de múltiplas camadas, consideremos a seguinte analogia:

- Dois engenheiros em países diferentes desejam trocar informações sobre um projeto de engenharia. Um engenheiro só fala português e o outro só se comunica em inglês. Para se comunicarem eles decidem utilizar um tradutor;
- Considere ainda, que o idioma comum entre os tradutores seja o alemão e que o meio utilizado para transmissão dos dados seja o telégrafo;
- Assim, o engenheiro que fala português passa suas informações para seu tradutor que as traduz para o alemão. A mensagem em alemão é então passada ao telegrafista que as transmite para um telegrafista no outro país;
- Ao receber a mensagem, o telegrafista passa para o tradutor que a traduz para o inglês e a entrega para o engenheiro. A figura abaixo ilustra essa comunicação, identificando os componentes da Arquitetura de Rede utilizada.

Nota-se que existe uma interface entre cada par de camadas adjacentes. É ela que definirá quais e como as funções oferecidas pela camada inferior podem ser acessadas pela camada superior. Esta interface deve ser bastante clara, de modo que, ao trocar-se a implementação de uma camada por outra completamente diferente, não seja necessário modificar as outras camadas. Isso é possível desde que a interface entre as camadas seja mantida. Por exemplo, trocando-se linhas telefônicas por transmissão via satélite, a implementação da camada responsável por manipular o acesso ao meio de transmissão deverá modificar completamente sua implementação, porém as demais camadas não sofrerão estas modificações desde que os mesmos serviços anteriores e o modo como são oferecidos sejam mantidos. Neste contexto o conjunto das camadas e protocolos é chamado de **ARQUITETURA DE REDE**.

8.2 ORGANIZAÇÕES INTERNACIONAIS DE PADRONIZAÇÃO

As primeiras arquiteturas de rede foram desenvolvidas por fabricantes de equipamentos, os quais desenvolviam soluções para interconexão apenas de seus produtos, sem se preocuparem com a compatibilidade de comunicação com equipamentos de outros fabricantes. Assim o fizeram, por exemplo, a IBM (*International Business Machines Corporation*) ao anunciar sua arquitetura de rede SNA (*System Network Architecture*), e a DEC (*Digital Equipament Corporation*) com sua DNA (*Digital Network Architecture*). Essas arquiteturas são denominadas **proprietárias**.

Desse modo, computadores de fabricantes diferentes não podiam se comunicar, impondo uma grande limitação aos consumidores, pois ficam “amarrados” aos produtos de um único fabricante, caso queira que seus equipamentos se comuniquem.

Torna-se evidente a necessidade de um conjunto de regras que permitam a comunicação ou interconexão entre dois sistemas quaisquer, sem considerar seu fabricante. Surgem as arquiteturas para interconexão de sistemas abertos: a **Arquitetura Internet**, desenvolvida por pesquisadores patrocinados pelo Departamento de Defesa dos Estados Unidos, e a **Arquitetura OSI** (*Open Systems Interconnection*) desenvolvida pela comunidade internacional sob a coordenação da ISO (**International Standards Organization**).

8.3 O MODELO DE REFERÊNCIA ISO RM-OSI

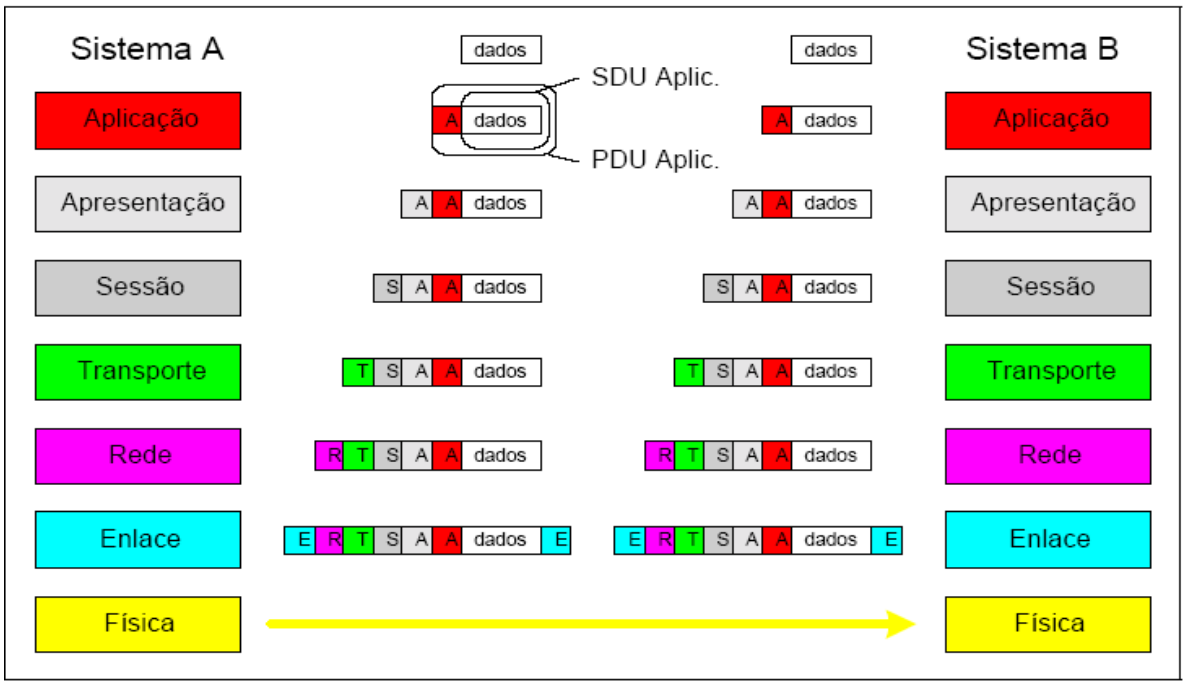
Baseada nas experiências advindas do funcionamento dos sistemas de teleprocessamento, da ARPAnet e das redes públicas e proprietárias, a ISO, entre 1978 e 1984, elaborou o "Modelo de Referência para Interconexão de Sistemas Abertos" (RM-OSI, *Reference Model for Open Systems Interconnection*), o qual define todos os princípios básicos para o desenvolvimento de uma arquitetura aberta.

O Modelo OSI por si só não é uma arquitetura de rede, pois não especifica exatamente os serviços e protocolos a serem usados em cada camada. Ele define alguns conceitos e divide a tarefa de comunicação em sete camadas funcionais, dizendo que funções cada camada deve desempenhar.

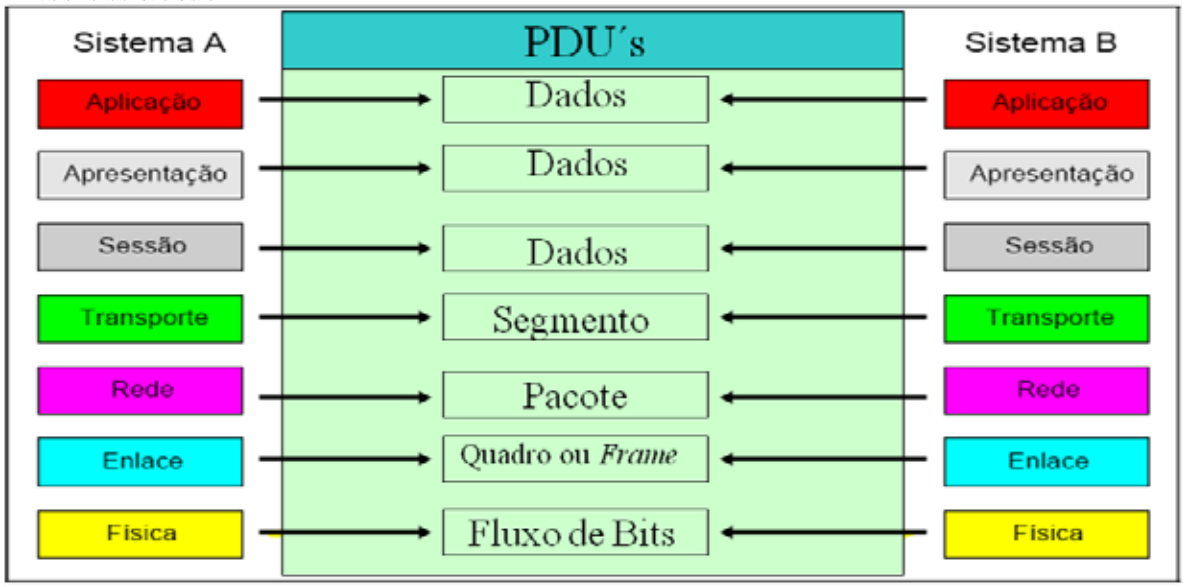
Entretanto, após elaborar o Modelo OSI, a ISO passou a projetar, especificar, implementar e testar os protocolos das várias camadas definidas pelo Modelo OSI, dando origem a **Arquitetura OSI**.

Neste curso nos limitaremos a citar somente a estrutura de camadas e as respectivas funções de cada camada como definido pelo Modelo OSI, sem entrar em detalhes dos protocolos de cada camada. As sete camadas do Modelo OSI estão representadas na figura ao lado.

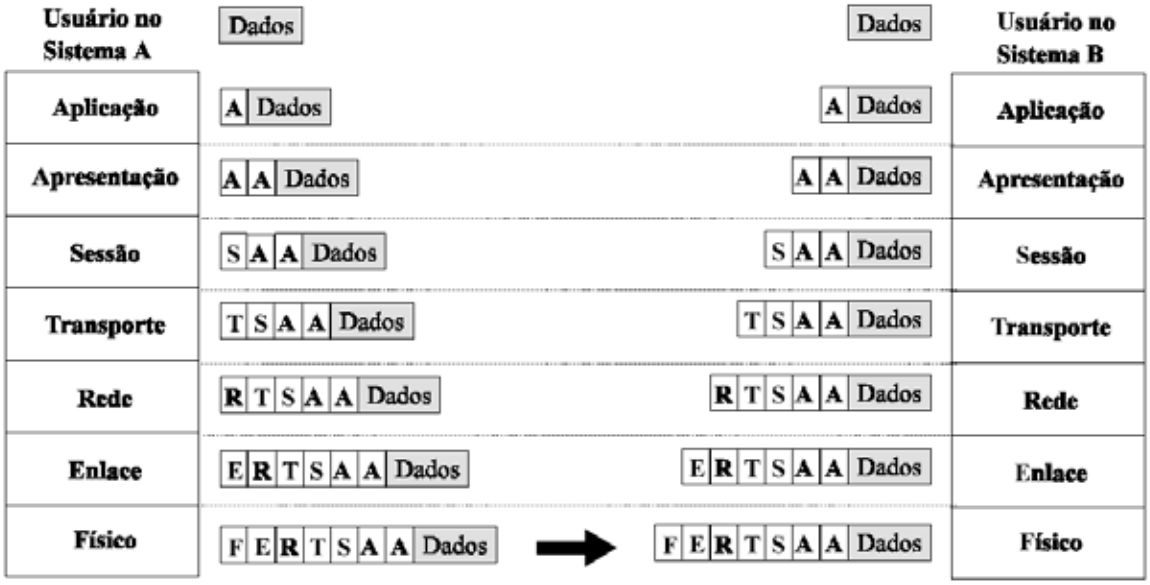
Embora o modelo OSI da ISO possa ser usado tanto em redes de longa distância quanto em redes locais, ele foi, em princípio, pensado para o uso em redes de longa distância. O Modelo OSI da ISO opera através das camadas na forma de encapsulamento, ou seja, em cada camada o dado será acrescido de um cabeçalho, posteriormente envelopados essas duas informações, o dado será enviado para a camada seguinte, onde será acrescido do cabeçalho seguinte, novamente envelopado, e assim sucessivamente até chegar na última camada.



Ao envelope, contendo o dado mais o cabeçalho da camada, denominamos de PDU (*Protocol Data Unit*, unidade de dados de protocolo). Quando o PDU passa para a camada seguinte, ele se transforma em dado na nova camada, esse novo dado (que por sua vez é composto de um dado bruto mais um cabeçalho) é denominado de SDU (*Service Data Unit*, unidade de dados de serviço). É importante saber que cada camada do RM-OSI possui seu PDU específico, sendo: “Dados” para as camadas de aplicação, apresentação e sessão; “Segmento” para a camada de transporte; “Pacote” para a camada de rede. “Quadro ou *Frame*” para a camada de enlace; E “Fluxo de Bits” para a camada física. Estes PDU também são considerados como as palavras chaves de cada camada, pois correspondem ao núcleo de maior importância.



A figura abaixo mostra como ocorre a transmissão de dados quando um usuário em um sistema A envia uma mensagem para um usuário em um sistema B, segundo o modelo OSI.



O processo começa com a entrega dos dados a serem transmitidos pelo usuário para a camada de aplicação na máquina A. A camada de aplicação junta aos dados do usuário um cabeçalho (*header*) contendo informações de controle de protocolo. Após isso, os dados do usuário, juntamente com o header anexado pela camada de aplicação são enviados para a camada de Apresentação. Para que possa executar sua função, esta também anexa suas informações de controle de protocolo e repassa os dados para a camada abaixo, ou seja, a camada de Seção. Esse processo é feito na máquina A até que cada camada faça sua função, ou seja, anexe seus *headers* de controle. Ao atingir a camada física na máquina A, os dados são transmitidos pelo meio de transmissão, juntamente com os *headers* colocados pelas camadas.

Na máquina B, ocorre o processo inverso. À medida que os dados vão sendo passados para as camadas superiores, cada camada retira o header colocado por sua camada correspondente na máquina origem (máquina A), executa as operações do protocolo de acordo com as informações contidas no *header,* e passa o restante para a camada superior. O processo se encerra com o usuário no sistema B recebendo os dados enviados pelo usuário do sistema A.

Não é intenção do padrão OSI-ISO servir como especificação de implementação, ou ser a base para conformar implementações já existentes, ou fornecer um nível de detalhes suficiente para a definição precisa dos serviços e protocolos da arquitetura proposta. O padrão fornece um esquema conceitual que permite que equipes de especialistas trabalhem de forma produtiva e independente no desenvolvimento de padrões para cada camada do RM-OSI.

O fato de dois sistemas distintos seguirem o RM-OSI não garante que eles possam trocar informações entre si, pois o modelo permite que sejam usadas diferentes opções de serviços/protocolos para as várias camadas. Essa flexibilidade pode levar a situações onde dois sistemas que utilizam opções de serviços/protocolos em conformidade com o RM-ISO não conseguem se comunicar, porque as opções adotadas são incompatíveis.

Para que dois sistemas quaisquer possam trocar informações é necessário que escolham opções compatíveis de serviço/protocolo para todas as camadas do modelo. Com o objetivo de definir grupos de opções de serviços/protocolos padronizados, a ISO elaborou o conceito de perfis funcionais. Se dois sistemas seguirem o mesmo perfil funcional eles com certeza irão comunicar-se, pois nesse caso as opções de serviço/protocolo adotadas serão compatíveis.

Pelo visto anteriormente, a coexistência de redes heterogêneas, fez com que se tornasse necessário definir uma arquitetura voltada para a interconexão dessas redes. Uma arquitetura importante no contexto de interconexão de redes heterogêneas é a Arquitetura desenvolvida para a Internet, que se baseia na família de protocolos **TCP/IP**, e que por apresentar características próprias, será descrita e comparada com o modelo OSI.

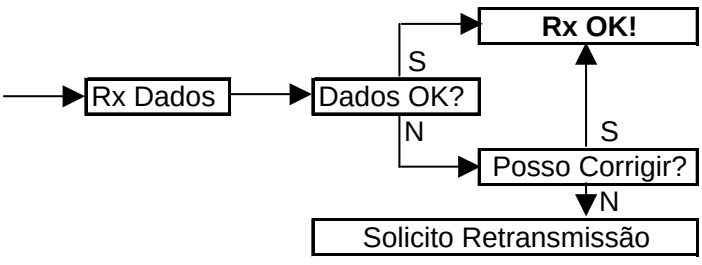
8.3.1 A Camada Física

A função desta camada é lidar com a transmissão pura de uma cadeia de bits através de um canal de comunicação. Deve garantir que, quando um lado transmite uma cadeia de 8 bits “1”, este seja recebido como 8 bits “1” do outro lado, e não com um bit “0” dentro da cadeia. Portanto, o protocolo da camada física deve considerar questões como: voltagem para bit "1"; voltagem para bit "0“; tempo de duração de um bit; o modo de transmissão (simplex, half-duplex, full-duplex); como a conexão é estabelecida e encerrada; pinagem dos conectores e etc. Ou seja, questões mecânicas, elétricas e funcionais da transmissão dos bits. Porém, o tratamento de erros de transmissão não é discutido neste nível.

8.3.2 A Camada de Enlace dos Dados

A principal função desta camada é detectar e, opcionalmente, corrigir possíveis erros que possam ocorrer durante a transmissão sobre o meio físico. O nível de enlace vai assim converter um canal de transmissão não confiável em um canal confiável para o uso do nível de rede. Para isso, ela particiona os dados recebidos da camada de rede em quadros (*frames*), algumas centenas de bits a serem enviados ao nível físico, adicionando cabeçalhos (*headers)* criando uma forma de redundância para detecção e controle de erros. Também deve cuidar da retransmissão de frames danificados ou perdidos e resolver problemas de duplicação de frames. Por exemplo, um ruído no meio de transmissão pode destruir o frame sendo transmitido. Neste Caso, a camada de Enlace de Dados na máquina origem deve retransmitir o frame. Entretanto, múltiplas retransmissões introduzem a possibilidade de frames duplicados.

Em geral, quase todos os protocolos de nível de enlace incluem bits de redundância em seus quadros para detecção de erros, mas não a sua correção. Essa técnica tem sido usada tradicionalmente devido ao fato da detecção e retransmissão requererem menos bits de redundância do que a correção. Com o crescente uso de transmissão via satélite (que possui um altíssimo retardo de transmissão), a



correção de erros se torna mais atrativa, já que dispensa o reenvio da informação. Cabe ressaltar que a função de correção de erros, quer por bits de redundância quer por retransmissão, é opcional neste nível de protocolo.

Outras tarefas desta camada são processar avisos de confirmação de recebimento enviados pelo receptor e resolver problemas de conexão entre máquinas com velocidades diferentes (quando uma máquina transmite dados em uma velocidade maior do que a máquina destino pode suportar, ocorrerá um estouro de *buffer* na máquina destino e os dados podem ser perdidos), dessa forma é utilizado algum mecanismo de controle de fluxo que possibilite o transmissor saber qual é o espaço disponível no *buffer* do receptor em um dado momento.

8.3.3 A Camada de Rede

O objetivo do nível de rede é fornecer ao nível de transporte uma independência quanto a considerações de chaveamento e roteamento associadas ao estabelecimento e operação de uma conexão de rede.

O nível de rede é responsável pela colocação da informação na Rede. Este nível verifica e envia as mensagens (ou pacotes) utilizando o endereço do nó de destino (host). Esta operação de enviar as mensagens (pacotes) está relacionada com o roteamento, que consiste no processo de procurar o menor e/ou melhor caminho para que esta mensagem (pacote) chegue ao seu endereço de destino. Este roteamento é baseado em cálculos que permitem determinar este caminho “ideal”. Chegando ao destino a mensagem (pacote) é entregue ao nível de transporte para processamento.

Em redes do tipo multi-ponto, mas com uma única rota (caso das redes LAN), devido à existência de um único canal, a função principal desse nível torna-se pouco relevante. Esse nível, nesse caso, poderá ser utilizado para permitir a interconexão entre redes (endereçamentos distinto) ou entre *sub-redes* (endereçamentos semelhantes).

Existem duas filosofias quanto ao serviço oferecido pelo nível de redes: Datagrama e Circuito virtual.

No serviço de **datagrama** (serviço não-orientado à conexão), cada pacote (unidade de dados do nível 3) não tem relação alguma de passado ou futuro com qualquer outro pacote, devendo assim carregar, de uma forma completa, seu endereço de destino. Nesse tipo de serviço, o roteamento é calculado toda vez que um pacote tem que ser encaminhado por um nó da rede.

No serviço de **circuito virtual** (serviço orientado à conexão), é necessário que o transmissor primeiramente envie um pacote de estabelecimento de conexão. A cada estabelecimento é dado um número, correspondente ao circuito, para uso pelos pacotes subseqüentes com o mesmo destino. Nesse método, os pacotes pertencentes a uma única conversação não são independentes.

Outras tarefas da camada de rede são: controle de congestionamento e tráfego; estatística de uso por usuário; resolver problemas de incompatibilidades (ex.: formas de endereçamento; tamanho de pacotes de dados; protocolos diferentes, etc) que podem ocorrer quando um pacote "viaja" por várias redes até alcançar a máquina destino.

8.3.4 A Camada de Transporte

O nível de rede não garante necessariamente que um pacote chegue a seu destino, e pacotes podem ser perdidos ou mesmo chegar fora da seqüência original de transmissão. Para fornecer uma comunicação fim a fim verdadeiramente confiável é necessário outro nível de protocolo, que é justamente o nível de transporte. Esse nível vai isolar dos níveis superiores a parte de transmissão da rede.

No nível de transporte, a comunicação é fim a fim, isto é, a entidade do nível de transporte da máquina de origem se comunica com a entidade do nível de transporte da máquina de destino. Isto pode não acontecer nos níveis físico, de enlace e de rede, onde a comunicação se dá muitas vezes entre máquinas adjacentes (vizinhas) na rede apenas para conexão entre a origem e o destino.

Duas funções importantes desse nível são a *multiplexação* (várias conexões de transporte compartilhando a mesma conexão de rede) e o *splitting* (uma conexão de transporte ligada a várias conexões de rede) de conexões. O *splitting* é utilizado para aumentar a vazão de uma conexão de transporte através do uso de várias conexões de rede simultaneamente. Já a *multiplexação* é usada quando uma conexão de transporte não gera tráfego suficiente para ocupar toda a capacidade da conexão de rede por ela utilizada.

Outra função importante do nível de transporte é o controle de fluxo. Como nenhuma implementação tem um espaço de armazenamento infinito, algum mecanismo deve ser fornecido de modo a evitar que o transmissor envie mensagens numa taxa maior do que a capacidade que o receptor tem de recebê-las. Além das funções mencionadas, pode-se ainda citar como funções desse nível o controle de seqüência fim a fim, a detecção e recuperação de erros fim a fim, a segmentação e blocagem de mensagens, o isolamento as camadas superiores das mudanças inevitáveis na tecnologia de hardware, entre outras.

8.3.5 A Camada de Sessão

Sua tarefa é permitir que usuários em máquinas diferentes estabeleçam sessões entre eles. Uma sessão permite a um usuário, por exemplo, realizar um *login* em um sistema de tempo compartilhado remoto ou transferir um arquivo entre duas máquinas. Esta camada é responsável por resolver todos os problemas que possam ocorrer durante uma sessão. Por exemplo:

• **controle de diálogo**: quando somente um lado da conexão pode transmitir em um dado instante (half-duplex), um mecanismo de *tokens* pode ser usado pela camada de sessão para esse fim;

• **sincronização da comunicação**: coloca pontos de checagem (sincronização) que permitem, em caso de quebra da comunicação, o restabelecimento da comunicação a partir do último ponto de sincronização checado. Ex. Transferência de arquivos.

Embora um circuito que permita transmissões nos dois sentidos seja necessário para o intercâmbio de informações, em algumas aplicações a troca de informações é half-duplex, ao invés de ser full-duplex. Com o intuito de fornecer um serviço de intercâmbio de informações half-duplex em um circuito full-duplex, o serviço de sessão utiliza o conceito de token. Em uma comunicação half-duplex. só o proprietário do token

de dados pode transmitir seus dados. O serviço de sessão fornece mecanismos para gerenciar a posse e passagem do token entre as entidades de aplicação que estão utilizando o serviço.

Em algumas aplicações, um volume muito grande de dados, por exemplo um arquivo extenso, é transmitido em redes muitas vezes não muito confiáveis. Embora o nível de transporte tente oferecer um circuito confiável, a rede pode simplesmente deixar de funcionar. Quando isso acontece, só resta ao nível de transporte indicar a falha e deixar a aplicação decidir o que deve ser feito. Eventualmente, a rede pode voltar a funcionar, podendo a conexão ser restabelecida. Nesse caso, o ideal seria que a transferência dos dados pudesse ser retomada do ponto imediatamente anterior ao da interrupção.

Com o objetivo de oferecer esse tipo de serviço, o nível de sessão usa o conceito de ponto de sincronização. Um ponto de sincronização é uma marca lógica posicionada ao longo do diálogo entre dois usuários do serviço de sessão. Toda vez que recebe um ponto de sincronização, o usuário do serviço de sessão deve responder com um aviso de recebimento ao usuário com quem está dialogando. Se por algum motivo a conexão for interrompida e depois restabelecida, os usuários podem retomar o diálogo a partir do último ponto de sincronização confirmado.

O conceito de atividade torna possível aos usuários do serviço de sessão distinguir partes do intercâmbio de dados, denominadas atividades. Cada atividade pode consistir em uma ou mais unidades de diálogo. Em uma conexão de sessão só é permitida a execução de uma atividade por vez, porém, podem existir várias atividades consecutivas durante a conexão. Uma atividade pode ser interrompida e depois recomeçada na mesma sessão, ou em conexões de sessão subseqüentes.

Para exemplificar o uso do conceito de atividade, consideremos o envio de uma mensagem através de um sistema de correio eletrônico como sendo uma atividade. Suponhamos que uma mensagem grande e de baixa prioridade esteja sendo transmitida e, durante a transmissão, a entidade do nível de sessão que a está enviando receba uma solicitação para enviar uma outra mensagem de maior prioridade. A entidade pode, então, suspender a atividade corrente, transferir a mensagem com alta prioridade, começando nesse caso uma outra atividade, e posteriormente retomar a atividade inicial (transmissão da mensagem de baixa prioridade).

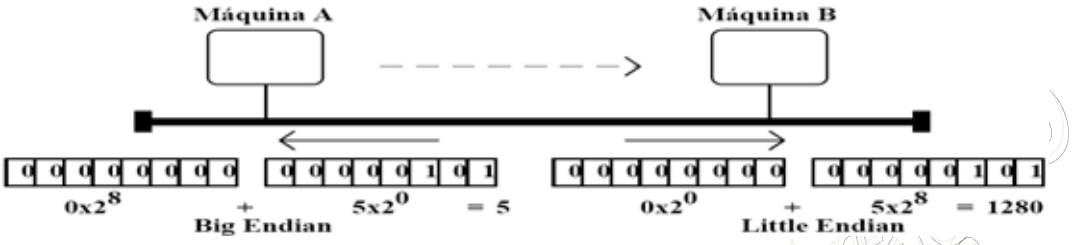
Utilizando o conceito de atividade, o nível de sessão permite também que dois usuários suspendam um diálogo, por exemplo no fim do expediente (obviamente desfazendo a conexão de sessão), e o retomem posteriormente, por exemplo no início do próximo expediente, utilizando uma nova conexão de sessão.

8.3.6 A Camada de Apresentação

Ao contrário das demais camadas que estão preocupadas em transferir dados de maneira confiável, a camada de Apresentação cuida da semântica e sintaxe da informação transferida. Ela permite que dados representando uma cadeia de caracteres, números reais ou inteiros, ou estrutura de dados, cheguem à máquina destino com o mesmo significado semântico e sintático com que foram transmitidos, independentemente dos diferentes padrões de codificação utilizados pelas máquinas envolvidas na comunicação. Isto é possível através da definição de dados em um modo abstrato, os quais podem ser convertidos para a representação padrão da rede. Por exemplo, suponhamos que a máquina A vai transmitir o número inteiro de dois bytes com valor 5 para a máquina B. Suponha ainda que a máquina A

utilize uma representação ***Big Endian***, onde o byte mais significativo é o da esquerda, e a máquina B utilize uma representação ***Little Endian***, com byte mais significativo à direita.

Ao receber o dado e transformá-lo na sua representação, a máquina B entenderia o número como tendo o valor 1280, ao invés de 5 como foi transmitido pela máquina A. Este problema é ilustrado na figura abaixo.



Dessa forma, a função do nível de apresentação é a de realizar transformações adequadas nos dados, antes de seu envio ao nível de sessão. Transformações típicas dizem respeito à compressão de textos, criptografia, conversão de padrões de terminais e arquivos para padrões de rede e vice-versa.

O nível de apresentação deve conhecer a sintaxe de seu sistema local bem como a sintaxe do sistema de transferência. Os serviços oferecidos por este nível são: transformação de dados, formatação de dados, seleção de sintaxes e estabelecimento e manutenção de conexões de apresentação.

8.3.7 A Camada de Aplicação

Fornece o suporte necessário para interação (comunicação) entre aplicações distribuídas, formando a interface entre um processo de usuário e os protocolos de comunicação. Nela estão serviços que são comumente necessários, tais como correio eletrônico, transferência de arquivos remotos, login remoto, etc.

Nesse nível são definidas funções de gerenciamento e mecanismos genéricos que servem de suporte à construção de aplicações distribuídas. Por exemplo, em grande parte das aplicações, para que seja possível o intercâmbio de informações é necessário estabelecer uma associação entre um ou mais usuários. Para realizar essa tarefa, o usuário do nível de aplicação pode utilizar um elemento de serviço da camada de aplicação denominado ACSE (*Association Control Service Element*).

Outros exemplos de elementos de serviço genéricos são o ROSE (*Remote Operations Service Element*) que oferece o suporte as chamadas de procedimentos remotos, e o RTSE (*Reliable Transfer Service Element*) que fornece um serviço de transferência de dados confiável, tornando todos os mecanismos de recuperação de erros transparentes aos usuários do serviço.

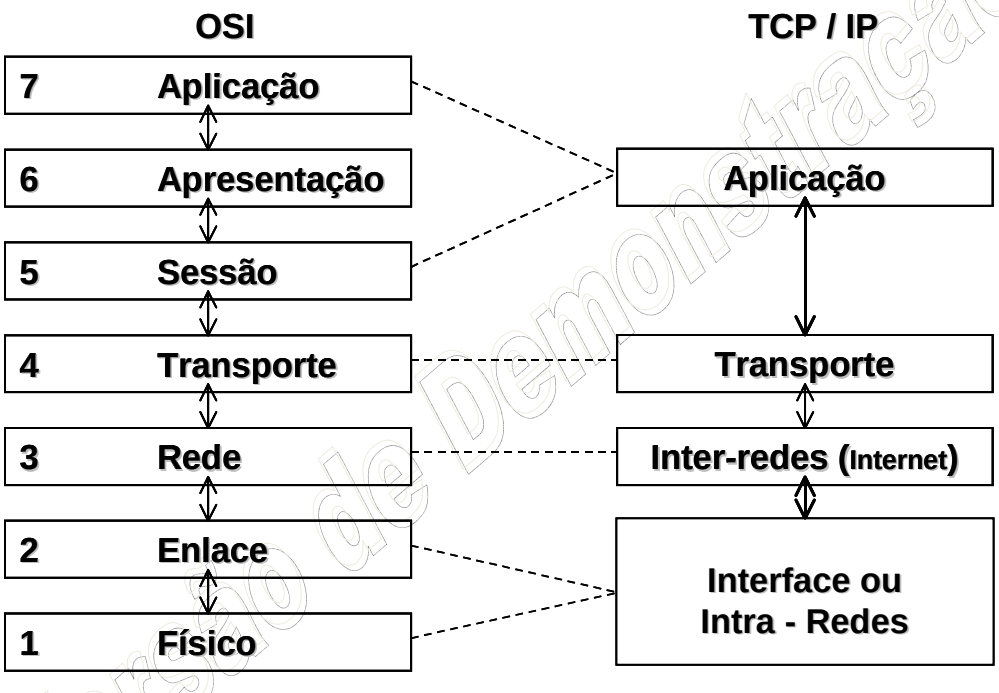
Além dos elementos de serviço genéricos, que são compartilhados pela maioria das aplicações, existem os elementos de serviço específicos de cada protocolo de aplicação, como o FTAM (*File Transfer, Access and Management)*, o DS (*Directory Service*), e o MHS (*Message Handling System*).

Na arquitetura TCP/IP, que veremos mais adiante, os responsáveis por esta camada são os programas utilizados na Internet, como o "protocolo de aplicação HTTP" para o navegador de Internet, o "protocolo de aplicação FTP" para a transferência de arquivos, e o "protocolo de aplicação SMTP e POP" para o envio e recebimento de e-mails, respectivamente.

9 COMPETÊNCIA 4 – FAMÍLIA DE PROTOCOLOS TCP/IP

9.1 COMPARAÇÕES COM O MODELO DE REFERÊNCIA RM-OSI/ISO

Como pode ser observado na figura abaixo, a primeira diferença entre as arquiteturas OSI e Internet TCP/IP está no número de camadas. Enquanto na arquitetura OSI são definidas sete camadas, na arquitetura TCP/IP são definidas quatro.

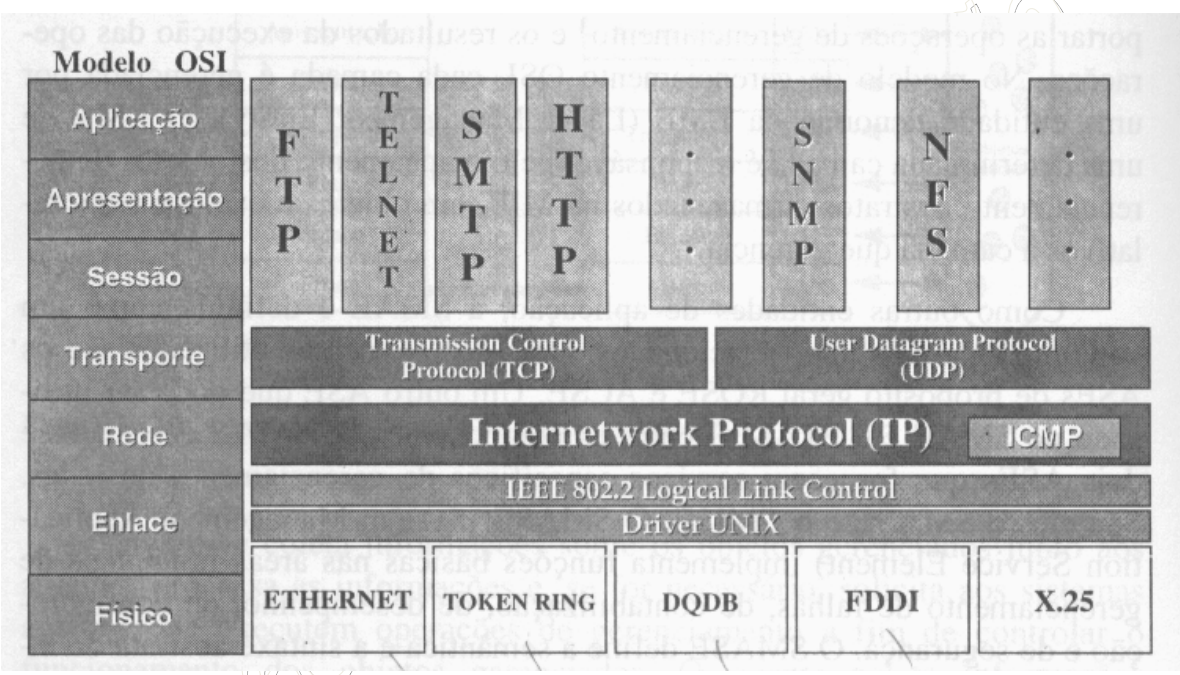


No RM-OSI são descritos formalmente os serviços de cada camada, a interface usada pelas camadas adjacentes para troca de informações e o protocolo que define regras de comunicação para cada uma das camadas. Alguns dos serviços definidos para as camadas do RM-OSI são opcionais. Por exemplo, os níveis de enlace, rede e transporte podem oferecer serviços orientados à conexão (circuito virtual) ou não-orientados à conexão (datagrama).

Essa característica é conseqüência do fato da ISO ter elaborado um modelo que se propõe a tratar todos os aspectos do problema de interconexão aberta de sistemas. Essa flexibilidade tem aspectos positivos, mas, por outro lado, pode levar a situações onde dois sistemas em conformidade com a arquitetura OSI não consigam se comunicar, bastando para tal que implementem perfis funcionais incompatíveis.

A arquitetura TCP/IP foi desenvolvida com o objetivo de resolver um problema prático: interligar redes com tecnologias distintas. Para tal, foi desenvolvido um conjunto especifico de protocolos que resolveu o problema de forma bastante simples e satisfatória. Os níveis: físico, enlace, e os aspectos do nível de rede do RM-OSI, relativos à transmissão de dados em uma única rede, não são abordados na arquitetura TCP/IP, que agrupa todos esses serviços na camada intra-rede. A arquitetura TCP/IP se limita a definir uma interface entre o nível intra-rede e o nível inter-rede.

Os serviços do nível de rede OSI relativos à interconexão de redes distintas são implementados na arquitetura TCP/IP pelo protocolo IP. Em outras palavras, nessa arquitetura só existe uma opção de protocolo e serviço para esta subcamada do nível de rede: o protocolo IP, cujo serviço é datagrama não confiável. Esta inflexibilidade da arquitetura TCP/IP no nível inter-rede é uma das principais razões de seu sucesso. O fato de um sistema utilizar ou não o protocolo IP foi usado inclusive para distinguir os sistemas que “estão na Internet” dos que não estão.



No nível de transporte, a arquitetura TCP/IP oferece duas opções: o TCP (que oferece um serviço de circuito virtual) e o UDP (datagrama). Esses protocolos são equivalentes aos protocolos orientados e não-orientados à conexão do nível de transporte OSI. Acima do nível de transporte está a camada de aplicações na arquitetura TCP/IP. Nessa arquitetura, os serviços dos níveis de sessão e apresentação OSI são implementados em cada aplicação de modo específico.

A abordagem da ISO, definindo as camadas de sessão, apresentação e elementos de serviços genéricos básicos no nível de aplicação, é mais razoável, no sentido em que permite uma maior reutilização de esforços durante o desenvolvimento de aplicações distribuídas.

Os protocolos da arquitetura TCP/IP oferecem uma solução simples, porém bastante funcional, para o problema da interconexão de sistemas abertos. O fato de implementações de seus protocolos terem sido a primeira opção de solução não-proprietária para a interconexão de sistemas fez com que essa arquitetura se tornasse um padrão de facto. A estrutura organizacional da ISO, com membros representando vários países, se por um lado aumenta o tempo de desenvolvimento dos padrões, por outro confere aos mesmos uma representatividade bem maior. Os padrões da ISO, por serem elaborados por uma instituição legalmente constituída para tal, são padrões de jure.

9.2 HISTÓRICO

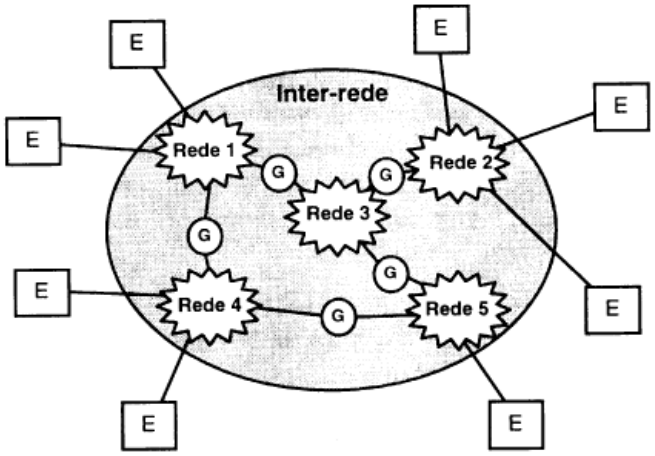
O desenvolvimento da arquitetura TCP/IP foi patrocinado pela *Defense Advanced Research Projects Agency* (DARPA), agência norte americana de pesquisas avançadas em defesa. A arquitetura baseia-se principalmente em: um serviço de transporte orientado à conexão, fornecido pelo *Transmission Control Protocol* (TCP), e em um serviço de rede não-orientado à conexão (datagrama não confiável), fornecido pelo Internet Protocol (IP).

Os padrões da arquitetura TCP/IP não são elaborados por órgãos internacionais de padronização, como a ISO ou o IEEE. O corpo técnico que coordena o desenvolvimento dos protocolos dessa arquitetura é um comitê denominado IETF (*Internet Engineering Task Force*). O IETF é formado por pesquisadores seniores, representantes de diversos governos, representantes de grandes empresas de telecomunicação, fornecedores de equipamentos, e usuários de Internet, tendo a maioria deles projetado e implementado os protocolos da Arquitetura Internet. O IETF, na realidade, produz poucos documentos. Qualquer pessoa pode projetar, documentar, implementar e testar um protocolo para ser usado na Internet.

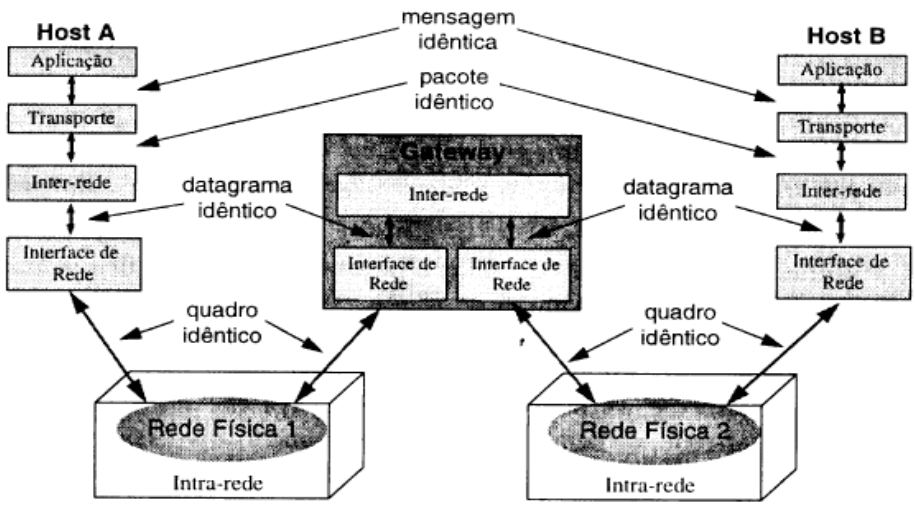
Para que um protocolo se torne um padrão Internet é necessário documentá-lo através de uma RFC (*Request for Comments*), http://www.ietf.org/rfc.html. As RFCs podem ser obtidas por qualquer pessoa conectada à Internet. Da análise das RFCs surgem sugestões, e novas versões do protocolo podem ser elaboradas. Quando o protocolo se torna estável, um dos membros do IAB propõe ao comitê que o protocolo se torne um padrão. Uma RFC é publicada modificando esse status e, se após decorridos aproximadamente seis meses não houver nenhuma objeção, o IETF declara o protocolo como um Internet Standard.

A arquitetura TCP/IP dá uma ênfase toda especial à interligação de diferentes tecnologias de redes. A idéia baseia-se na seguinte constatação: não existe nenhuma tecnologia de rede que atenda aos anseios de toda a comunidade de usuários. Alguns usuários precisam de redes de alta velocidade que normalmente cobrem uma área geográfica restrita. Já outros, se contentam com redes de baixa velocidade que conectam equipamentos distantes milhares de quilômetros uns dos outros. Portanto, a única forma de permitir que um grande volume de usuários possa trocar informações é interligar as redes às quais eles estão conectados, formando assim uma inter-rede.

Para interligar duas redes distintas é necessário conectar uma máquina a ambas as redes. Tal máquina fica responsável pela tarefa de transferir mensagens de uma rede para a outra. Uma máquina que conecta duas ou mais redes é denominada *internet gateway* ou *internet router*. Para ser capaz de rotear corretamente as mensagens, os gateways precisam conhecer a topologia da inter-rede, ou seja, precisam saber como as diversas redes estão interconectadas. Já os usuários vêem a inter-rede como uma rede virtual única à qual todas as máquinas estão conectadas, não importando a forma física de interconexão, como exemplifica ao lado.



A arquitetura internet TCP/IP é organizada em quatro camadas conceituais construídas sobre uma quinta camada que não faz parte do modelo, a camada intra-rede.



No nível de **aplicação,** os usuários usam programas de aplicação para acessar os serviços disponíveis na inter-rede. As aplicações interagem com o nível de **transporte** para enviar e receber dados. As aplicações podem usar o serviço orientado à conexão, fornecido pelo TCP (serviço de circuito virtual), ou o serviço não-orientado à conexão, fornecido pelo *User Datagram Protocol* — UDP (serviço de datagrama não confiável).

Qualquer comunicação precisa de padrões para que as partes comunicantes se entendam. Da mesma forma acontece com os computadores: ao se comunicarem, os computadores precisam trocar dados dentro de uma padronização conhecida por ambos, caso contrário não irá ser efetuada esta troca.

Existem várias "línguas computacionais", na verdade dentro de uma arquitetura geral para as quais foram criadas. Para o usuário não técnico as mesmas receberam a designação geral de “protocolo”, cada um com suas peculiaridades (não esquecer o que no vocabulário técnico o termo protocolo se refere aos padrões de comunicação entre os níveis de uma arquitetura ...). O “protocolo” TCP/IP pode ser considerado, que é mais ou menos o "inglês da Internet", ou seja, de utilização geral.

Antes da popularização da Internet, existiam diferentes protocolos sendo utilizados nas redes das empresas. Os mais utilizados eram TCP/IP; NETBEUI; IPX/SPX; Apple Talk. À medida que a Internet tornou-se mais popular, com o aumento exponencial do número de usuários, o protocolo TCP/IP passou a ser um padrão de fato, utilizado não só na Internet, mas também em redes internas das empresas que começavam a ser conectadas à Internet.

Como as redes internas precisavam conectar-se à Internet, tinham que usar o mesmo protocolo da Internet, ou seja: TCP/IP. Dos principais sistemas operacionais do mercado, o Unix sempre utilizou o protocolo TCP/IP como padrão. O Windows dá suporte ao protocolo TCP/IP desde as primeiras versões, porém o TCP/IP somente tornou-se o protocolo padrão a partir do Windows 2000.

Ser protocolo padrão significa que o TCP/IP será instalado durante a instalação do sistema operacional, a não ser que um protocolo diferente seja selecionado. Até mesmo o sistema operacional Novell, que sempre foi baseado no IPX/SPX, passou a adotar o TCP/IP como padrão a partir da versão 5.0.

O que temos hoje, na prática, é a utilização do protocolo TCP/IP na esmagadora maioria das redes e sua adoção é cada vez maior. Como não poderia deixar de ser, o TCP/IP é o protocolo padrão do Windows 2000 e também do Windows XP. Se durante a instalação, o Windows detectar a presença de uma placa de rede, automaticamente será sugerida a instalação do protocolo TCP/IP.

Revisando o **Capítulo 1**, devemos relembrar "como" os computadores vão se interligar para trocar informações. Resumidamente, existem dois modos básicos para isto: a comutação de circuitos e a comutação de pacotes. Na comutação de circuitos, os computadores se ligam diretamente para a troca de informações, e na comutação de pacotes, os computadores mandam a informação para pontos intermediários até alcançarem seu destino.

Enquanto na comutação de circuitos os computadores podem trocar grandes volumes de informações, de forma direta e contínua, na comutação de pacotes, os computadores mandam pequenos pedaços (pacotes) de informações que, por não irem diretamente ao destino, precisam conter - em cada pacote - o endereço de origem e destino completos. Comparando com nosso cotidiano, tomemos os exemplos do telefone e da carta: no telefone, estabelecemos um circuito e falamos diretamente com o destinatário, entretanto, só podemos falar com uma pessoa de cada vez. Com as cartas é diferente, pois além de podermos mandar cartas para várias pessoas ao mesmo tempo, não temos controle sobre os caminhos que estas cartas percorrerão até seus destinos. Nem mesmo ao mandarmos várias cartas para a mesma pessoa! Elas podem tomar rumos diferentes a cada envio...

Este dois tipos de comunicação foram analisados na criação do protocolo TCP/IP e, de acordo com as necessidades da época, uma se mostrou mais útil que a outra. Quais eram as necessidades? Primeiro, as redes a serem construídas não poderiam parar, mesmo que alguns dos computadores desta rede fossem destruídos; Segundo, os computadores não poderiam estar, todos, fisicamente ligados (por motivos de custo); Terceiro, esta rede possuiria vários tipos de computadores diferentes, e eles precisariam se comunicar.

Não foi preciso muito esforço para perceber que uma comutação de circuitos seria cara e frágil, por causa da falta de ligações redundantes (secundárias), o que seria extremamente útil e fácil de implementar em uma comutação de pacotes. Portanto, foi criado um protocolo que se comunica por pacotes: o TCP/IP.

Este protocolo foi criado em "camadas" ou níveis, ou seja, dento do próprio protocolo, existem processos bem definidos, cada um fazendo sua tarefa na comunicação, e cada processo se comunicando com o seguinte através dos níveis adjacentes. O modelo de camadas tem uma vantagem óbvia: se os processos de comunicação estão bem definidos e separados em cada camada, qualquer alteração necessária em um destes processos poderia ser feita isoladamente, não precisando, portanto, reescrever todo o protocolo.

9.3 NÍVEL FÍSICO

Conforme vimos na última seção, a camada física do modelo OSI/ISO é aglutinada na camada de intra-rede. Porém, alguns estudiosos e profissionais consideram a existência desta quinta camada como uma forma de referenciar a parte física da mídia de comunicação, de bits, de quadros, de endereçamento MAC, etc. Em nosso curso, não iremos considerar esta quinta camada, e sim que ela está aglutinada na camada intra-rede do TCP/IP.

9.4 NÍVEL DE INTRA-REDES E INTERFACES DE REDES

Este nível, dependendo do meio ao qual está ligado, enviará um pacote diferente para cada caso. Por exemplo, se for uma placa Ethernet, enviará os quadros padrão IEEE 802.3, se for ATM, Frame Relay, Token Ring, enviará seus quadros específicos. O importante nesta camada, para o funcionamento do TCP/IP, não é a mesma em si, mas sim a maneira com que a camada superior se comunica com ela

Como exemplos e representantes diretos dessa camada, encontramos a padronização 802.2, muito importante e extremamente utilizada, que foi desenvolvida por iniciativa do *Institute of Electrical and Electronics Engineers* (IEEE), e que define os níveis físicos e enlace de redes locais de computadores.

OSI está para aspectos lógicos como IEEE 802 está para aspectos físicos da rede

A família 802 de padrões agrupa:

- Cabos
- Adaptadores de redes
- Transceptores
- Conectores de cabos
- Dispositivos para conexão (hubs, switchs, etc)

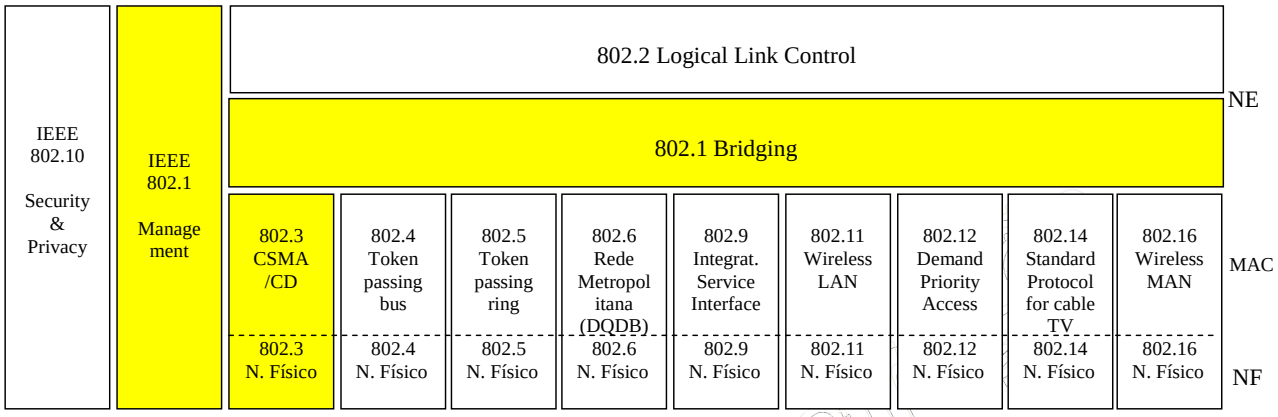
A família 802 de padrões limita:

- Velocidades de transmissões
- Acesso a redes
- Distância de cabos
- Concentrações de dispositivos

Os principais representantes da família IEEE 802, e que estão diretamente associados ao nível de intra-redes do TCP/IP, são:

802.2	Logical Link Control (LLC)	802.10	Security & Privacy access
802.3	CSMA/CD (Ethernet)	802.11	Wireless LAN access (WiFi)
802.4	LAN Token-Passing Bus	802.12	Demand-Priority Access Method
802.5	LAN Token-Passin Ring	802.14	Standard Protocol for cable TV
802.6	MAN (DQDB)	802.15	Wireless PAN (Blue Tooth)
802.7	Conselho Técnico de BandaLarga	802.16	Rede Metropolitana Wireless (WiMax)
802.8	Conselho Técnico de Fibra Ótica	802.17	Metropolitan Area Network (RPR - Resilient Packet Ring)
802.9	Rede de Integração de Voz e Dados	802.20	Broadband Wireless Access

Todos esses padrões estão diretamente ou indiretamente relacionados para compor essa família, veja o diagrama abaixo para compreender este relacionamento:



Entre esses padrões, o que iremos estudar com maiores detalhe são os protocolos de acesso ao meio, em especial o 802.3, os protocolos de passagem de permissão 802.4 e 802.5, e os protocolos de redes sem fio 802.11 e 802.16.

9.4.1 Protocolos de Acesso Múltiplos ao Meio

Existem dois tipos de enlaces de redes: enlaces ponto-a-ponto e enlaces *broadcast*. Um enlace ponto-a-ponto consite em um único remetente em uma extremidade do enlace e um único receptor na outra extremidade do enlace. Muitos protocolos de camada de enlace foram projetados para enlaces ponto-a-ponto; o PPP (protocolo ponto-a-ponto) é um desses protocolos. O segundo tipo de enlace, o enlace *broadcast*, pode ter vários nós remetentes e receptores, todos conectados ao mesmo canal de transmissão único e compartilhado. O termo *broadcast* é usado aqui porque, quando qualquer um dos nós transmite um quadro, o canal propaga o quadro e cada um dos nós recebe uma cópia. A Ethernet e as LANs sem fio são exemplos de tecnologias de *broadcast* de camada de enlace.

O problema dos enlaces broadcast consiste em como coordenar o acesso de vários nós remetentes e receptores a um canal compartilhado. Como todos os nós têm a capacidade de transmitir quadros, mais do que dois nós podem transmitir ao mesmo tempo. Quando isso acontece, todos os nós recebem vários quadros ao mesmo tempo, isto é, os quadros transmitidios colidem em todos os receptores. Em geral, quando há uma colisão, nenhum dos nós receptores consegue perceber algum sentido nos quadros que foram transmitidos; de certo modo, os sinais dos quadros que colidem ficam inextricavelmente embaralhados. Assim, todos os quadros envolvidos na colisão são perdidos e o canal broadcast é desperdiçado durante o intervalo de colisão. É claro que, se muitos nós querem transmitir quadros freqüentemente, muitas transmissões resultarão em colisões e grande parte da largura de banda do canal broadcast será desperdiçada.

A solução é encontrada através dos protocolos de acesso múltiplo, pelos quais os nós regulam sua transmissão pelos canais compartilhados. Esses protocolos são classificados nas seguintes categorias: protocolos de divisão de canal, protocolos de acesso aleatório e protocolos de revezamento.

Os protocolos de divisão de canal permitem que um determinado nó (host) fale durante um período de tempo fixo, em seguida, permite que um outro nó fale pelo mesmo período de tempo e assim por diante. Os principais protocolos são o TDM (Time Division Multiplex), o FDM (Frequency Division Multiplex) e o CDMA (Code Division Multiple Access). O TDM e FDM atribuem aos nós intervalos de tempo e frequência, respectivamente, evitando colisões e dividindo a largura de banda com justiça entre os nós. Porém, também compartilham de uma desvantagem, mesmo quando um único nó tem pacotes a enviar sua largura de banda é limitada e fixa. O protocolo CDMA atribui um código diferente a cada nó. Então, cada nó usa seu código exclusivo para codificar os bits de dados que envia. Se os códigos forem escolhidos cuidadosamente, as redes CDMA terão a propriedade de permitir que nós diferentes transmitam simultaneamente e, ainda assim, consigam que seus receptores respectivos recebam corretamente os bits codificados pelo remetente (admitindo-se que o receptor conheça o código do remetente), dessa forma o protocolo CDMA consegue evitar colisões, dividir a largura de banda com justiça entre os nós e não limitar a largura de banda.

O protocolo CDMA, devido as suas propriedades antiinterferências, é muito utilizado em redes militares e de canais de acesso múltiplo sem fio.

A segunda classe geral de protocolos de acesso múltiplo são os protocolos de acesso aleatório. Essa classe de protocolos permite que um nó transmissor sempre transmita à taxa total do canal. Quando há uma colisão, cada nó envolvido nela retransmite repetidamente seu quadro (pacote) até que este passe sem colisão. Mas, quando um nó sofre uma colisão, ele nem sempre retransmite o quadro imediatamente. Em vez disso, ele espera um tempo aleatório antes de retransmitir o quadro. Cada nó envolvido de uma colisão escolhe atrasos aleatórios independentes. Como após uma colisão os tempos de atraso são escolhidos independentemente, é possível que um dos nós escolha um atraso suficientemente mais curto do que os atrasos dos outros nós em colisão e, portante, consiga passar seu quadro discretamente para dentro do canal, sem colisão.

O primeiro dos protocolos dessa segunda classe é o ALOHA, ou slotted Aloha, apesar de não estar mais em uso, contudo seu histórico é importante, pois dele se deriva o Ethernet, o protocolo da atualidade. O ALOHA tem esse nome em função do seu criador ser um fã do surf, e ter homenageado a faculdade do Hawai. Diferentemente da partição de canal, esse protocolo permite que um único nó transmita continuamente à taxa total do canal, quando ele for o único nó ativo. Também apresenta a característica da descentralização, onde cada nó detecta colisões e decide independentemente quando retransmitir, adotando um mecanismo de sincronização dos intervalos nos nós.

O slotted Aloha funciona bem quando há somente um nó ativo, mas qual é sua eficiência quando há vários nós ativos? Nesse caso, há duas preocupações possíveis. A primeira é que, quando há vários nós ativos, uma certa fração dos intervalos terá colisões e, portanto, será desperdiçada. A segunda preocupação é que uma outra fração dos intervalos estará vazia porque todos os nós ativos evitarão transmitir como resultado da política probabilística de transmissão. Os únicos intervalos não desperdiçados serão aqueles em que exatamente um nó transmite. Um intervalo em que exatamente um nó transmite é denominado um intervalo bem-sucedido. A eficiência de um protocolo de acesso múltiplo com intervalos é definida como a fração (calculada durante um longo tempo) de intervalos bem-sucedidos no caso em que há um grande número de nós ativos, cada qual tendo sempre um grande número de quadros a enviar. Note que, se não fosse usado nenhum tipo de controle de acesso e cada nó retransmitisse imediatamente após cada colisão, a eficiência seria zero. O slotted Aloha claramente aumenta a eficiência para além de zero.

O protocolo slotted Aloha requer que todos os nós sincronizem suas transmissões para que comecem no início de um intervalo. O primeiro protocolo ALOHA era, na realidade, um protocolo sem intervalos e totalmente descentralizado. No ALOHA puro, quando um quadro chega pela primeira vez (isto é, um datagrama de camada de rede é passado para baixo a partir da camada de rede no nó remetente), o nó imediatamente transmite o quadro inteiro ao canal broadcast. Se um quadro transmitido sofrer uma colisão com uma ou mais transmissões, o nó retransmitirá imediatamente (após ter conluído a transmissão total do quadro que sofreu a colisão) o quadro com certa probabilidade. Caso contrário, o nó esperará por um tempo de transmissão de quadro. Após essa espera, ele então retransmitirá o quadro com mesma probabilidade ou espera (permanece ocioso) por um outro tempo de quadro com probabilidade menor.

Tanto no slotted Aloha quanto no ALOHA pura, a decisão de transmitir tomada por um nó independe da atividade dos outros nós ligados ao canal broadcast. Em particular, um nó não se preocupa se por acaso um outro nó está transmitindo quando ele começa a transmitir nem pára de transmitir se outro nó começar a interferir em sua transmissão. Essa ansiedade de transmissão é melhor controlada por mecanismos que ouvem o meio antes de falar, denominada de detecção de portadora, e pelo parar de falar se alguém começar a falar ao mesmo tempo que você, denominada de detecção de colisão.

Na detecção de portadora, um nó ouve o canal antes de transmitir. Se um quadro de outro nó estiver correntemente sendo transmitido para dentro do canal, o nó então esperará (se afastará – back off) por um período de tempo aleatório e, então, novamente sondará o canal. Se perceber que o canal está ocioso, o nó então começará a transmitir quadros. Caso contrário, ele esperará por um outro período aleatório de tempo e repetirá esse processo.

Na detecção de colisão, um nó que está transmitindo ouve o canal enquanto transmite. Se esse nó detectar que outro nó está transmitindo um quadro interferente, ele pára de transmitir e usa algum protocolo para determinar quando deve tentar transmitir novamente.

Essas duas regras estão incorporadas na família de protocolos CSMA (Carrier Sense Multiple Access – acesso múltiplo com detecção de portadora) e CSMA/CD (CSMA com detecção de colisão).

Lembre-se de que duas propriedades desejáveis de um protocolo de acesso múltiplo são: (1) quando apenas um nó está ativo, esse nó ativo tem uma vazão de R bps; (2) quando M nós estão ativos, então cada nó ativo tem uma vazão de aproximadamente R/M bps. Os protocolos ALOHA e CSMA têm a primeira propriedade, mas não a segunda. Isso motivou os pesquisadores a criarem uma outra classe de protocolos, os protoclos de revezamento. Como acontece com os protocolos de acesso aleatório, há dezenas de protocolos de revezamento, e cada um desses protocolos tem muitas variações. Os mais importantes são o protocolo de seleção (polling) e o protocolo de passagem de permissão (token).

O protocolo de polling requer que um dos nós seja designado como nó mestre. O nó mestre seleciona cada um dos nós por alternância circular. Em particular, ele envia primeiramente uma mensagem ao nó 1 dizendo que ele (o nó 1) pode transmitir até um certo número máximo de quadros. Após o nó 1 transmitir alguns quadros, o nó mestre diz ao nó 2 que ele (o nó 2) pode transmitir até um certo número máximo de quadros. (O nó mestre pode terminar quando um nó terminou de enviar seus quadros observando a ausência de um sinal no canal.) O procedimento continua dessa maneira, com o nó mestre escolhendo cada um dos nós de maneira cíclica.

O protocolo de polling elimina as colisões e os intervalos vazios que atormentam os protocolos de acesso aleatório, o que permite que ele tenha uma eficiência muito maior. Mas esse protocolo também de algumas desvantagens. A primeira desvantagem é que o protocolo introduz um atraso de seleção – o período de tempo requerido para notificar um nó que ele pode transmitir. A segunda desvantagem é potencialmente mais séria: se o nó mestre falhar, o canal inteiro ficará inoperante.

No protocolo de passagem de permissão não há nó mestre. Um pequeno quadro de finalidade especial conhecido como uma permissão (token) é passado entre os nós obedecendo a uma determinada ordem fixa. Por exemplo, o nó 1 poderá sempre enviar a permissão ao nó 2, o nó 2 poderá sempre enviar a permissão ao nó 3, o nó N poderá sempre enviar a permissão ao nó 1. Quando um nó recebe uma permissão, ele a retém somente se tiver alguns quadros para transferir, caso contrário, imediatamente a repassa para o nó seguinte. Se um nó tiver quadros para transmitir quando recebe a permissão, ele enviará um número máximo de quadros e, em seguida, passará a permissão para o nó seguinte. A passagem de permissão é descentralizada e tem uma alta eficiência. Mas também tem seus problemas. Por exemplo, a falha de um nó pode derrubar o canal inteiro. Ou, se um nó acidentalmente se descuida e não libera a permissão, então é preciso chamar algum procedimento de recuperação para colocar a permissão novamente em circulação. Exemplos desses protocolos são o FDDI e o IEEE 802.5.

9.4.2 Passagem de Permissão

Protocolos de acesso múltiplo são usados em conjunto com muitos tipos diferentes de canais broadcast. Eles têm sido utilizados por canais de satélite e sem fio, cujos nós transmitem sobre um espectro de freqüência comum. Atualmente eles estão sendo usados no canal de acesso por cabo à Internet na direção usuário-provedor e são utilizados extensivamente em redes locais (LANs).

Na década de 1980 e no início da década de 1990, duas classes de tecnologias de LAN eram populares nos ambientes de trabalho. A primeira classe consistia nas LANs Ethernet (também conhecidas como LANs 802.3), que eram redes de acesso aleatório. A segunda classe de tecnologias de LAN compreendiam as tecnologias de passagem de permisão, incluindo a token ring (também conhecida como IEEE 802.5) e a FDDI – interface de dados distribuída de fibra. Nossa discussão sobre as tecnologias de passagem de permissão é intencionalmente breve, já que a inexorável concorrência da Ethernet praticamente as extinguiu. Mesmo assim, para dar exemplos de tecnologia de passagem de permissão e apresentar uma pequena perspectiva histórica, será útil falar um pouco sobre anéis de passagem de permissão, ou token rings.

Em uma LAN token ring, os N nós da LAN (hospedeiros e roteadores) estão conectados em um anel por enlaces diretos. A topologia do anel define a ordem de passagem de permissão. Quando um nó obtém a permissão e envia um quadro, este se propaga ao redor do anel interior, criando, dessa maneira, um canal virtual de transmissão broadcast. À medida que o quadro se propagada, o nó de destino lê esse quadro no meio de transmissão da camada de enlace. O nó que envia o quadro tem a responsabilidade de remover o quadro do anel. A FDDI foi projetada para LANs de alcance geográfico maior, incluindo as denominadas redes de área metropolitana (MAN). Para LANs de grande alcance geográfico (WAN), é ineficiente permitir que um quadro se propague de volta ao nó remetente tão logo tenha passado do nó de destino. A FDDI faz com que o nó de destino remova o quadro do círculo. (Estritamente falando, a FDDI não é um canal broadcast puro, pois todos os nós recebem todos os quadros transmititdos).

9.4.3 Padrão IEEE 802.3 (CSMA/CD – Ethernet)

A Ethernet praticamente tomou conta do mercado de LANs com fio. Na década de 1980 e início da década de 1990, ela enfrentou muitos desafios de outras tecnologias LAN, incluindo token ring, FDDI e ATM. Algumas dessas outras tecnologias conseguiram conquistar uma parte do mercado de LANs durante alguns anos. Mas, desde sua invenção, em meados da década de 1970, a Ethernet continuou a se desenvolver e crescer e conservou sua posição dominante no mercado. Hoje, ela é de longe a tecnologia preponderante de LAN com fio e é provável que continue assim no futuro próximo. Podemos dizer que a Ethernet está sendo para a rede local o que a Internet tem sido para a rede global.

Há muitas razões para o sucesso da Ethernet. Em primeiro lugar, ela foi a primeira LAN de alta velocidade amplamente disseminada. Como foi disponibilizada cedo, os administradores de rede ficaram bastante familiarizados com a Ethernet, e relutaram em mudar para outras tecnologias LAN quando estas apareceram em cena. Em segundo lugar, token ring, FDDI e ATM são tecnologias mais complexas e mais caras do que a Ethernet, o que desencorajou ainda mais os administradores na questão da mudança. Em terceiro lugar, a razão mais atraente para mudar para outra tecnologia LAN (como FDDI e ATM) era normalmente a velocidade mais alta da nova tecnologia; contudo, a Ethernet sempre se defendeu produzindo versões que funcionavam a velocidades iguais, ou mais altas. E, também, a Ethernet comutada foi introduzida no início da década de 1990, o que aumentou ainda mais sua velocidade efetiva de dados. Finalmente, como a Ethernet se tornou muito popular, o hardware para Ethernet (em particular, adaptadores hubs e concentradores) se tornou mercadoria comum, de custo muito baixo.

A LAN Ethernet original foi inventada em meados da década de 1970 por Bob Metcalfe e David Boggs. Usava um barramento para interconectar os nós, que persistiu durante toda a década de 1980 e por grande parte da década de 1990; em particular, a tecnologia Ethernet 10Base2, que utilizava um cabo coaxial fino para o barramento, era imensamente popular na década de 1990. Contudo, exceto uma ocasional instalação herdada, quase todas as instalações Ethernet de hoje utilizam uma topologia em estrela, com um hub ou comutador no centro.

Todas as tecnologias Ethernet fornecem serviço não orientado para conexão à camada de rede. Isto é, quando o adaptador A quer enviar um datagrama ao adaptador B, o adaptador A encapsula o datagrama em um quadro Ethernet e envia o quadro à LAN, sem se conectar previamente a B. Esse serviço de camada 2 não orientado para conexão é análogo ao serviço de datagrama de camada 3 do IP e ao serviço de camada 4 não orientado para conexão do UDP.

Todas as tecnologias Ethernet fornecem um serviço não confiável à camada de rede. Especificamente, quando o adaptador B recebe um quadro do adaptador A, ele submete o quadro a uma verificação de CRC, mas não envia um reconhecimento quando um quadro passa na verificação CRC nem um reconhecimento negativo quando um quadro não passa na verificação de CRC. Quando um quadro não passa na verificação de CRC, o adaptador B simplesmente o descarta. Assim, o adaptador A não têm a mínima idéia se o quadro que transmitiu passou na verificação CRC. Essa falta de transporte confiável (na camada de enlace) ajuda a tornar a Ethernet simples e barata. Mas também significa que a seqüência de datagramas passada à camada de rede pode ter lacunas, e que essas camadas superiores precisarão tratar da confiança da entrega.

Quando os nós estão interconectados com um hub (e não a um comutador de camada de enlace) a LAN Ethernet é uma verdadeira LAN de broadcast, isto é, quando um adaptador transmite um quadro, todos os adaptadores na LAN recebem o quadro. Como pode empregar broadcast, a Ethernet precisa de um protocolo de acesso múltiplo – ela usa o CSMA/CD, que resumidamente realiza:

1. Um adaptador pode começar a transmitir a qualquer tempo, ou seja, não são usados compartimentos.
2. Um adaptador nunca transmite um quadro quando percebe que algum outro adaptador está transmitindo, ou seja, ele usa detecção de portadora.

3. Um adaptador que está transmitindo aborta sua transmissão quando percebe que algum outro adaptador está transmitindo, ou seja, usa detecção de colisão.
4. Antes de tentar uma retransmissão, um adaptador espera um período de tempo aleatório que é caracteristicamente pequeno em comparação com o tempo de transmissão de um quadro.

Esses mecanismos conferem ao CSMA/CD um desempenho muito melhor, em ambientes LAN, do que o do slotted Aloha, De fato, se o atraso máximo de propagação entre estações for muito pequeno, a eficiência do CSMA/CD poderá ficar próxima a 100 por cento. Mas note que o segundo e o terceiro mecanismo que citamos requerem que cada adaptador Ethernet seja capaz de (1) perceber quando algum outro adaptador está transmitindo e (2) detectar uma colisão enquanto estiver transmitindo. Adaptadores Ethernet realizam essas duas tarefas medindo os níveis de tensão antes e durante a transmissão.

9.4.4 Tecnologias Ethernet

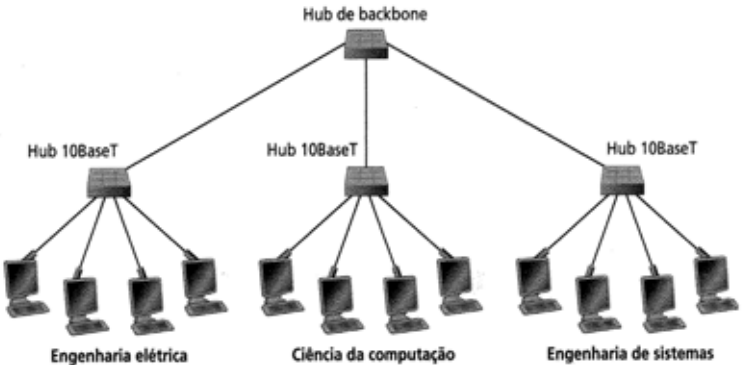
Em 2004, as tecnologias Ethernet mais comuns eram a 10BaseT e a 100BaseT, que usam pares de fios trançados de cobre em topologia estrela e têm velocidades de transmissão de 10 Mbps e 100 Mbps, respectivamente. Essas tecnologias Ethernet estão padronizadas pelos grupos de trabalho IEEE 802.3. Por essa razão, uma LAN Ethernet quase sempre é denominada LAN 802.3.

O “T” em 10BaseT e em 100BaseT quer dizer “par de fios trançados”. Tanto para o 10BaseT quanto para o 100BaseT, o comprimento máximo da conexão entre um adaptador e o hub é 100 metros; assim, a distância máxima entre quaisquer dois nós é 200 metros.

Um repetidor (hub) é um dispositivo de camada física que atua sobre bits individuais, em vez de sobre quadros, e tem duas ou mais interfaces. Quando um bit, representando um 0 ou um 1, chega de uma interface, o repetidor simplesmente recria o bit, reforça sua energia e transmite o bit para todas as outras interfaces. É importante ter em mente que repetidores não implementam detecção de portadora ou qualquer outra parte do CSMA/CD; um repetidor repete um bit que está entrando em todas as interfaces de saída mesmo que haja energia de sinal em algumas dessas interfaces. Como repetidores fazer transmissão broadband de bits, cada adaptador em uma Ethernet 10/100 BaseT pode (1) sondar o canal para determinar se ele está ocioso e (2) detectar uma colisão enquanto está transmitindo.

9.4.5 Hubs, Comutadores e Roteadores

O modo mais simples de interconectar as LANs é utilizar hubs.



A figura ao lado mostra como três departamentos de uma universidade podem interconectar suas LANs. Nessa figura, cada um dos três departamentos tem uma Ethernet 10BaseT que fornece acesso à rede ao corpo acadêmico, ao pessoal e aos estudantes do departamento.

Cada hospedeiro em um departamento tem uma conexão ponto-a-ponto com o hub do departamento. Um quarto hub, denominado hub de backbone, tem conexões ponto-a-ponto com os hubs dos departamentos, interconectando as LANs dos três departamentos. O desenho mostrado é um projeto de hub multinível, porque os hubs são arranjados hierarquicamente. Também é possível criar projetos multiníveis com mais de dois níveis – por exemplo, um nível para os departamentos, um nível para as escolas dentro da universidade (escola de engenharia, escola de administração ,etc) e um nível mais alto para o ambiente universitário.

Em um projeto multinível, referimo-nos à rede interconectada como uma LAN e a cada uma das parcelas departamentais da LAN (isto é, a cada hub departamental e aos hospedeiros conectados a ele) como um segmento de LAN. É importante notar que todos os segmentos de LAN pertencem ao mesmo domínio de colisão, isto é, sempre que um ou mais nós nos segmentos de LAN transmitem ao mesmo tempo, há colisão e todos os nós transmissões entram em backoff.

Uma LAN departamental interconectada a um hub de backbone tem muitos benefícios. Em primeiro lugar, e mais importante, fornece comunicação interdepartamental entre os hospedeiros dos vários departamentos. Em segundo lugar, amplia a distância máxima entre qualquer par de nós da LAN. Por exemplo, com uma 10BaseT, a distância máxima entre um nó e seu hub é 100 metros; portanto, para um único segmento de LAN, a distância máxima entre qualquer par de nós é 200 metros. Interconectando os hubs, essa distância máxima pode ser aumentada, já que a distância entre hubs diretamente conectados também pode ser de 100 metros quando se usa par trançado (e maior ainda quando se usa fibra). Um terceiro benefício é que o projeto multinível provê um certo grau de degradação suave. Especificamente, se qualquer um dos hubs departamentais começar a funcionar mal, o hub de backbone poderá detectar o problema e desconectar o hub departamental da LAN; desse modo, os departamentos restantes podem continuar a funcionar e a se comunicar enquanto o hub departamental defeituoso é consertado.

Embora um hub de backbone seja um dispositivo de interconexão útil, ele tem três sérias limitações que dificultam sua disseminação. A primeira limitação, e talvez a mais importante, é que quando LANs departamentais estão interconectadas por um hub, os domínios de colisão dos departamentos (anteriormente independentes) se transforma em um grande domínio de colisão comum. Antes de interconectar os três departamentos, cada LAN departamental tinha uma vazão máxima de 10 Mbps, de modo que a vazão agregada máxima das três LANs era de 30 Mbos. Mas, assim que as três LANs são interconectadas por um hub, todos os hospedeiros dos três departamentos passam a pertencer ao mesmo domínio de colisão e a vazão agregada máxima fica reduzida a 10 Mbsps.

Uma segunda limitação é que, se os vários departamentos usarem tecnologias Ethernet diferentes, então poderá ser impossível interconectar os hubs departamentais a um hub de backbone. Por exemplo, se alguns departamentos usam 10BaseT e os departamentos restantes usam 100BaseT, é impossível interconectar todos os departamentos sem fazer algum buffer de quadros no ponto de interconexão; uma vez que um hub é, essencialmente, um repetidor que não armazena quadros, ele não pode interconectar segmentos de LAN que funcionam em velocidades diferentes.

Uma terceira limitação é que cada uma das tecnologias Ethernet (10Base2, 10BaseT, 100BaseT e assim por diante) tem restrições quanto ao número máximo de nós permitidos em um domínio de colisão, quanto à distância máxima à distância máxima entre dois nós dentro de um domínio de colisão e quanto ao número máximo de níveis permitidos em um projeto multinível. Essas restrições limitam tanto o número total de hospedeiros que podem se conectar a uma LAN multinível quanto o alcance geográfico de uma LAN multinível.

Em contraste com os hubs, que são dispositivos de camada física, comutadores de camada de enlace – denominados simplesmente de comutadores (switchs) – agem sobre quadros Ethernet e, portanto, são dispositivos de camada 2. De fato, como são comutadores de pacotes totalmente capacitados, eles repassam quadros com base em endereços da LAN de destino. Quando um quadro chega à interface de um comutador, o comutador examina o endereço de destino de camada 2 do quadro e tenta repassá-lo para a interface que leva a esse destino.

Comutadores podem superar muitos problemas que atormentam os hubs. Primeiramente, eles permitem comunicação interdepartamental, preservando, ao mesmo tempo, domínios de colisão isolados para cada um dos segmentos LAN.

Em segundo lugar, eles podem interconectar diferentes tecnologias de LAN, incluindo as Ethernet 10BaseT, 100BaseT e a Gigabit Ethernet. Em terceiro lugar, não há limites ao tamanho possível de uma LAN quando são usados comutadores para interconectar segmentos de LAN; teoricamente, utilizando comutadores, é possível montar uma LAN que abranja o globo terrestre inteiro. E também, comutadores funcionam em full-duplex e fornecem comutação acelerada (CUT-though switching).

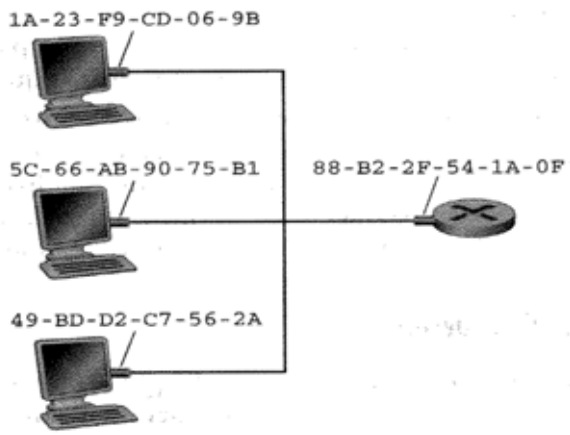
O modo full-duplex pode ser compreendido neste caso como uma conexão ponta-a-ponta entre um nó A e um nó B por meio de um comutador. Vamos supor que os nós A, B, C e D operam em 10BaseT, ou seja, vazão de 10 Mbps. Havendo uma transferência de dados entre A e B e entre C e D, implica dizer que a vazão máxima de dados entre A-B é de 10Mbps e entre C-D é de também 10Mbps. Ou seja, o comutador permite o isolamento das comunicações.

O modo de comutação de pacotes acelerada (cut-through) difere do modo de comutação de pacotes do tipo armazenagem e repasse (store-and-forward), usada por roteadores. Na comutação acelerada, um pacote não precisa ser completamente armazenado antes de ser repassado, em vez disso, ele é repassado através do comutador quando o enlace de saída está livre, reduzindo o atraso fim-a-fim. Se o enlace de saída é uma rede de acesso múltiplo compartilhada com outros hospedeiros (por exemplo, o enlace de saída se conecta com um hub), então o comutador também tem de sondar se o enlace está ocioso antes de iniciar a retransmissão dos pacotes por comutação acelerada.

Um roteador por sua vez é equivalente a um comutador de enlace, porém adota o modo de operação store-and-forward, onde os pacotes são primeiramente armazenados por completo até sua completa chegada, para em seguida serem retransmitidos. Além dessa característica, os roteadores operam sobre os endereços de pacotes de camada 2, enquanto os comutadores de enlace operam sobre os endereços de enlace MAC. Apesar de suas similaridades, entretanto suas aplicações são distintas. Os comutadores de enlace existem para unificar segmentos de LAN equivalentes, isolando os domínios de broadcast. Os roteadores existem para unificar segmentos de LAN disjuntos, ofertando o mecanismo de filtram por firewall entre as duas redes.

9.4.6 Endereçamento de Enlace

Na verdade, não é o nó (isto é, o hospedeiro ou o roteador) que tem um endereço de camada de enlace, mas o adaptador do nó.



Um endereço de camada de enlace é também denominado um endereço de LAN, um endereço físico ou um endereço MAC (*media access control* – controle de acesso ao meio). Como a expressão endereços MAC parece ser o mais popular, daqui em diante nos referiremos a endereços de camada de enlace como endereços MAC.

Para a maior parte das LANs (incluindo a Ethernet e as LANs 802.11 sem fio), o endereço MAC tem 6 bytes de comprimento, o que dá 2^{48} possíveis endereços MAC.

Como ilustrado na figura, esses endereços de 6 bytes são tipicamente expressos em notação hexadecimal, com cada byte do endereço expresso como um par de números hexadecimais. Um fato importante referente aos endereços MAC é que eles são permanentes – quando um adaptador é fabricado, um endereço MAC é gravado na ROM do adaptador.

Uma propriedade interessante dos endereços MAC é que não existem dois adaptadores com o mesmo endereço. Isso pode parecer surpreendente, dado que os adaptadores são fabricados em muitos países por inúmeras empresas diferentes. Como uma empresa fabricante de adaptadores em Taiwan se certifica de que está usando endereços diferentes dos usados por uma empresa fabricante de adaptadores na Bélgica? A resposta é que o IEEE gerencia o espaço físico de endereços MAC. Em particular, quando uma empresa quer fabricar adaptadores, compra, por uma taxa nominal, uma parcela do espaço de endereços que consiste em 2^{24} endereços. O IEEE aloca a parcela de 2^{24} endereços fixando os primeiros 24 bits de um endereço MAC e permitindo que a empresa crie combinações exclusivas com os últimos 24 bits para cada adaptador.

O endereço MAC de um adaptador tem uma estrutura linear (oposta à estrutura hierárquica) e nunca muda, não importando para onde vá o adaptador. Um computador portátil com um cartão Ethernet tem sempre o mesmo endereço MAC, não importando para onde o computador vá. Um PDA com uma interface 802.11 tem sempre o mesmo endereço MAC onde quer que vá. Lembre-se de que, ao contrário, um endereço IP tem uma estrutura hierárquica (isto é, uma parte que é da rede e uma parte que é do hospedeiro) e que o endereço IP de um nó precisa ser trocado quando o hospedeiro muda de lugar. O endereço MAC de um adaptador é análogo ao número do CPF de uma pessoa, que também tem uma estrutura linear e não muda, não importando para onde a pessoa vá. Um endereço IP é análogo ao endereço postal de uma pessoa, que é hierárquico e precisa ser trocado quando a pessoa muda de lugar. Exatamente como uma pessoa pode achar útil ter um endereço postal, bem como um número de CPF, também é útil para um nó ter um endereço de camada de rede, bem como em endereço MAC.

Como descrevemos no início desta seção, quando um adaptador quer enviar um quadro para algum adaptador de destino, o adaptador remetente insere no quadro o endereço MAC do destino e envia o quadro para dentro da LAN. Se a LAN utilizar transmissão broadcast (como a LAN 802.11 e muitas LANs Ethernets), o quadro será recebido e processado por todos os outros adaptadores na LAN. Em particular, cada adaptador que recebe o quadro verificará se o endereço MAC de destino que está no quadro combina com seu próprio endereço MAC. Se os endereços combinarem, o adaptador extrairá o datagrama encerrado no quadro e o passa para cima na pilha de protocolos até seu nó pai. Se os endereços não combinarem, o adaptador descartará o quadro sem passar o datagrama de camada de rede para cima na pilha de protocolos. Assim, somente o adaptador no nó de destino interromperá seu nó pai quando receber um quadro.

No entanto, às vezes um adaptador remetente quer que todos os outros adaptadores na LAN recebam e processem o quadro que ele está prestes a enviar. Nesse caso, o adaptador remetente isnere um endereço de broadcast MAC especial no campo de endereço do destinatário do quadro. Para LANs que usam endereços de 6 bytes (como a Ethernet e as LANs de passagem de permissão), o endereço de broadcast é uma cadeia de 48 bits 1 consecutivos (isto é, FF-FF-FF-FF-FF-FF em notação hexadecimal).

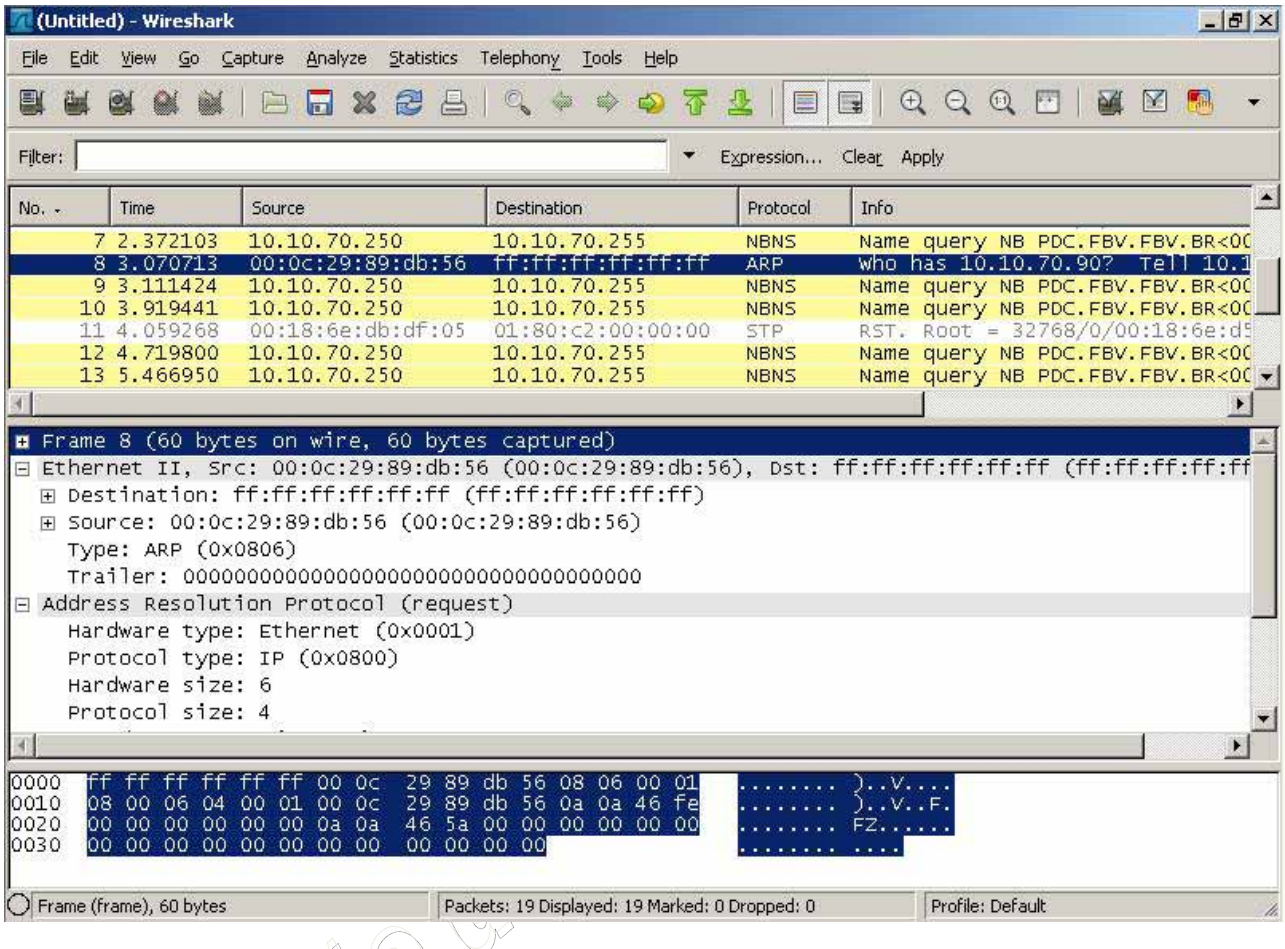
Como existem endereços de camada de rede (por exemplo, endereços IP da Internet) e endereços de camada de enlace (isto é, endereços MAC), é preciso fazer a tradução de um para o outro. Para a Internet, esta é uma tarefa do protocolo de resolução de endereços (*addess resolution protocol* – ARP) [RFC 826].

O ARP é uma função que recebe como entrada o endereço IP e retorna o endereço MAC correspondente na rede do endereço IP.

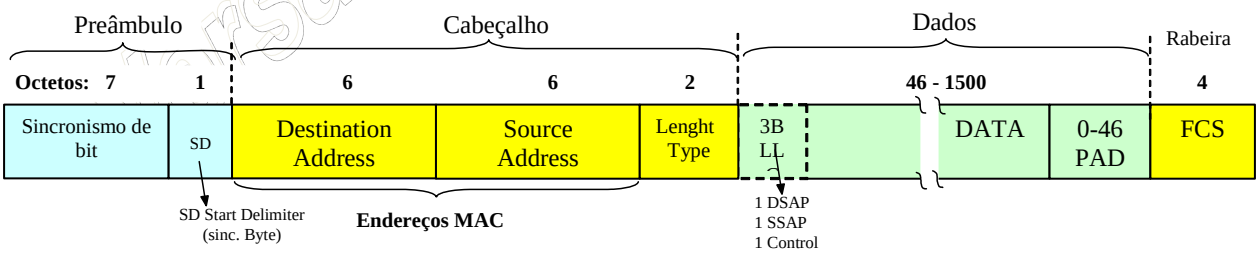
Endereço IP	Endereço MAC	TTL
222.222.222.221	88-B2-2F-54-1A-0F	13:45:00
222.222.222.223	5C-66-AB-90-75-B1	13:52:00

Cada nó (hospedeiro ou roteador) tem em sua RAM uma tabela ARP que contém mapeamentos de endereços IP para endereços MAC, além de um valor de tempo de vida (TTL) que indica quando cada mapeamento será apagado da tabela. Note que a tabela não contém necessariamente um registro para nó da sub-rede; alguns nós podem ter tido registros que já expiraram, ao passo que outros nós podem jamais ter sido registrados na tabela. Um tempo de remoção típico para um registro é de 20 minutos a partir do momento em que foi colocado em uma tabela ARP.

Quando um nó deseja se comunicar com outro nó, porém não possui na tabela ARP um registro mapeado, a estação transmissora, através do protocolo ARP, envia uma requisição ARP Query através de broadcast para o segmento de rede. O nó de destino, ao receber a requisição responde através do ARP Reply. Os nós, os comutadores e roteadores que estiverem entre esta comunicação, ao observarem o tráfego, irão adicionar a suas tabelas ARP as respectivas resoluções.



Quadro MAC do Padrão IEEE 802.3



9.5 NÍVEL DE INTER-REDES

Também chamada de "Interredes" ou “internet” (com “i” minúsculo) esta camada é a responsável pelo direcionamento do tráfego dos pacotes. É nela que são identificados os endereços IP (IP vem de "Internet Protocol"). Serve para as comunicações intermediárias entre um computador e outro. Esta camada possui também um protocolo para envio de mensagens de controle e de erro, que é enviado para as camadas "Interredes" de outros computadores, chamado ICMP (Internet Control Message Protocol). O ICMP, muito resumidamente, tem a função de, por exemplo, checar a existência de um endereço.

Um exemplo desse teste é o comando "ping" Ele manda pacotes ICMP do tipo "Tem alguém aí?" para o destino especificado e obtém uma resposta que diz alguma coisa tipo "Sim! Estou aqui!". Detalhe: os

pacotes ICMP são enviados dentro dos pacotes IP, assim como os da camada de transporte, mas o ICMP é gerado dentro na própria camada "Inter-redes".

9.5.1 Endereçamento IP

O endereçamento IP é sempre um tema importante, já que é ele que permite que o brutal número de redes e hosts que formam a Internet sejam capazes de se comunicar entre si.

Existem duas versões do protocolo IP: o **IPV4** é a versão atual, que utilizamos na grande maioria das situações, enquanto o **IPV6** é a versão atualizada, que prevê um número brutalmente maior de endereços e deve se popularizar a partir de 2012 ou 2014, quando os endereços IPV4 começarem a se esgotar.

No IPV4, os endereços IP são compostos por 4 blocos de 8 bits (32 bits no total), que são representados através de números de 0 a 255 (cobrindo as 256 possibilidades permitidas por 8 bits), como "200.156.23.43" ou "64.245.32.11". Os grupos de 8 bits que formam o endereço são chamados de "octetos", o que dá origem a expressões como "o primeiro octeto do endereço". De qualquer forma, a divisão dos endereços em octetos e o uso de números decimais serve apenas para facilitar a configuração para nós, seres humanos. Quando processados, os endereços são transformados em binários, como "11001000100110010001011100101011".

As faixas de endereços começadas com "10", "192.168" ou de "172.16" até "172.31" são reservadas para uso em redes locais e por isso não são usadas na Internet. Os roteadores que compõe a grande rede são configurados para ignorar pacotes provenientes destas faixas de endereços, de forma que as inúmeras redes locais que utilizam endereços na faixa "192.168.0.x" (por exemplo) podem conviver pacificamente, sem entrar em conflito.

No caso dos endereços válidos na Internet, as regras são mais estritas. A entidade global responsável pelo registro e atribuição dos endereços é a IANA (http://www.iana.org/), que delega faixas de endereços às RIRs (Regional Internet Registries), entidades menores, que ficam responsáveis por delegar os endereços regionalmente. Nos EUA, por exemplo, a entidade responsável é a ARIN (http://www.arin.net/) e no Brasil é a LACNIC (http://www.lacnic.net/pt/). Estas entidades são diferentes das responsáveis pelo registro de domínios, como o Registro.br.

As operadoras, carriers e provedores de acesso pagam uma taxa anual à RIR responsável, que varia de US$ 1.250 a US$ 18.000 (de acordo com o volume de endereços requisitados) e embutem o custo nos links revendidos aos clientes. Note que estes valores são apenas as taxas pelo uso dos endereços, não incluem o custo dos links, naturalmente.

Ao conectar via ADSL ou outra modalidade de acesso doméstico, você recebe um único IP válido. Ao alugar um servidor dedicado você recebe uma faixa com 5 ou mais endereços e, ao alugar um link empresarial você pode conseguir uma faixa de classe C inteira. Mas, de qualquer forma, os endereços são definidos "de cima para baixo" de acordo com o plano ou serviço contratado e você não pode escolher quais endereços utilizar.

Embora aparentem ser uma coisa só, os endereços IP incluem duas informações: o endereço da rede e o endereço do host dentro dela. Em uma rede doméstica, por exemplo, você poderia utilizar os endereços

"192.168.1.1", "192.168.1.2" e "192.168.1.3", onde o "192.168.1." é o endereço da rede (e por isso não muda) e o último número (1, 2 e 3) identifica os três micros que fazem parte dela.

Os micros da rede local podem acessar a Internet através de um roteador, que pode ser tanto um servidor com duas placas de rede quando um modem ADSL ou outro dispositivo que ofereça a opção de compartilhar a conexão. Nesse caso, o roteador passa a ser o gateway da rede e utiliza seu endereço IP válido para encaminhar as requisições feitas pelos micros da rede interna. Esse recurso é chamado de NAT (Network Address Translation).

Um dos micros da rede local, neste caso, poderia usar esta configuração de rede:

Endereço IP: 192.168.1.2
Máscara: 255.255.255.0
Gateway: 192.168.1.1 (o servidor compartilhando a conexão)
DNS: 200.169.126.15 (o DNS do provedor)

O servidor, por sua vez, utilizaria uma configuração similar a esta:

Placa de rede 1 (rede local):
Endereço IP: 192.168.1.1
Máscara: 255.255.255.0

Placa de rede 2 (Internet):
Endereço IP: 200.213.34.21
Máscara: 255.255.255.0
Gateway: 200.213.34.1 (o gateway do provedor)
DNS: 200.169.126.15 (o DNS do provedor)

A configuração da segunda placa de rede seria obtida automaticamente, via DHCP, de forma que você só precisaria realmente se preocupar com a configuração da sua rede local. Normalmente, você primeiro configuraria a rede local, depois conectaria o servidor à Internet e, depois de checar as duas coisas, ativaria o compartilhamento da conexão via NAT.

O servidor DHCP incluído no ICS do Windows utiliza uma configuração fixa, fornecendo endereços dentro da faixa "192.168.0.x", mas ao utilizar um servidor Linux, ou qualquer outro dispositivo de rede que ofereça um servidor DHCP com mais recursos, você pode escolher qualquer faixa de endereços e também configurar uma "zona" para os endereços do servidor DHCP, permitindo que você tenha micros com IPs fixos e IPs dinâmicos (fornecidos pelo servidor DHCP) na mesma rede. Nesse caso, você poderia ter uma configuração como a seguinte:

192.168.0.1: Gateway da rede
192.168.0.2: Ponto de acesso wireless
192.168.0.3: Servidor de arquivos para a rede interna
192.168.0.4 até 192.168.0.99: Micros da rede configurados com IP fixo
192.168.0.100 até 192.168.0.254: Faixa de endereços atribuída pelo servidor DHCP

Veja que usar uma das faixas de endereços reservadas não impede que os PCs da sua rede possam acessar a Internet. Embora eles não acessem diretamente, por não possuírem IPs válidos, eles podem acessar através de uma conexão compartilhada via NAT ou de um servidor proxy. É possível, inclusive, configurar o firewall ativo no gateway da rede para redirecionar portas (port forwarding) para micros dentro da rede local, de forma que eles possam ser acessados remotamente. O servidor nesse caso "empresta" uma porta, ou uma determinada faixa de portas, para o endereço especificado dentro da rede local. Quando

alguém da Internet acessa uma das portas encaminhadas no servidor, é automaticamente redirecionado para a porta correspondente no micro da rede interna, de forma transparente.

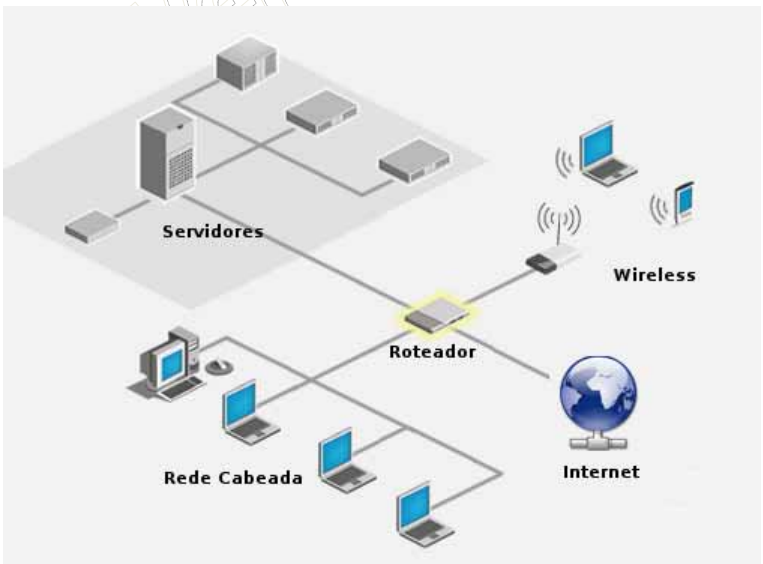
O uso dos endereços de rede local tem aliviado muito o problema da falta de endereços IP válidos, pois uma quantidade enorme de empresas e usuários domésticos, que originalmente precisariam de uma faixa de endereços completa para colocar todos os seus micros na Internet, pode sobreviver com um único IP válido (compartilhado via NAT entre todos os micros da rede). Em muitos casos, mesmo provedores de acesso chegam a vender conexões com endereços de rede interna nos planos mais baratos, como, por exemplo, alguns planos de acesso via rádio, onde um roteador com um IP válido distribui endereços de rede interna (conexão compartilhada) para os assinantes.

Embora seja possível, pelo menos em teoria, ter redes com até 24 milhões de PCs, usando a faixa de endereços de rede local 10.x.x.x, na prática é raro encontrar segmentos de rede com mais de 100 ou 200 micros. Conforme a rede cresce, o desempenho acaba caindo, pois, mesmo ao utilizar um switch, sempre são transmitidos alguns pacotes de broadcast (que são retransmitidos a todos os micros do segmento). A solução nesse caso é dividir a rede em segmentos separados, interligados por um roteador.

Em uma empresa, poderíamos (por exemplo) ter três segmentos diferentes, um para a rede cabeada (e a maior parte dos micros), outro para a rede wireless e outro para os servidores.

O roteador nesse caso teria 4 interfaces de rede (uma para cada um dos três segmentos e outra para a Internet). A vantagem de dividir a rede desta maneira é que você poderia criar regras de firewall no roteador, especificando regras diferentes para cada segmento. Os micros conectados à rede wireless (menos segura), poderiam não ter acesso aos servidores, por exemplo. Quando falo em "roteador", tenha em mente que você pode perfeitamente usar um servidor Linux com diversas placas de rede.

Com relação à proteção da rede contra acessos provenientes da Internet, você poderia tanto configurar o próprio firewall ativo no roteador, de forma a proteger os micros da rede local quanto instalar um firewall dedicado (que pode ser um PC com duas placas de rede) entre ele e a Internet:

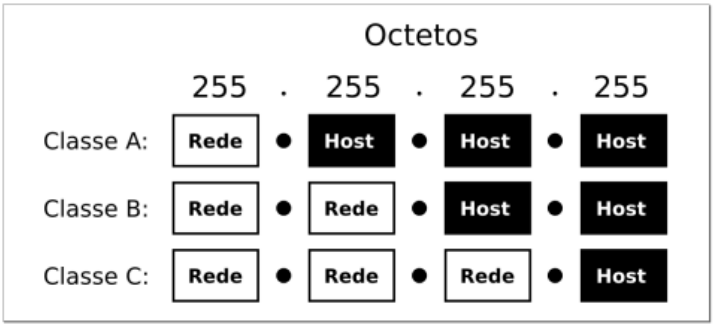


Voltando à questão dos endereços: inicialmente os endereços IP foram divididos em classes, denominadas A, B, C, D e E. Destas, apenas as classe A, B e C são realmente usadas, já que as classes D e E são reservadas para recursos experimentais e expansões futuras.

Cada classe reserva um número diferente de octetos para o endereçamento da rede. Na classe A, apenas o primeiro octeto identifica a rede, na classe B são usados os dois primeiros octetos e na classe C temos os três primeiros octetos reservados para a rede e apenas o último reservado para a identificação dos hosts dentro dela.

O que diferencia uma classe de endereços da outra é o valor do primeiro octeto. Se for um número entre 1 e 126 temos um endereço de classe A. Se o valor do primeiro octeto for um número entre 128 e 191,

então temos um endereço de classe B e, finalmente, caso o primeiro octeto seja um número entre 192 e 223, temos um endereço de classe C.



Ao configurar uma rede local, você pode escolher a classe de endereços mais adequada. Para uma pequena rede, uma faixa de endereços de classe C (como a tradicional 192.168.0.x com máscara 255.255.255.0) é mais apropriada, pois você precisa se preocupar em configurar apenas o último octeto do endereço ao atribuir os endereços. Em uma rede de maior porte, com mais de 254 micros, passa a ser necessário usar um endereço de classe B (com máscara 255.255.0.0), onde podemos usar diferentes combinações de números nos dois últimos octetos, permitindo um total de 65.534 endereços.

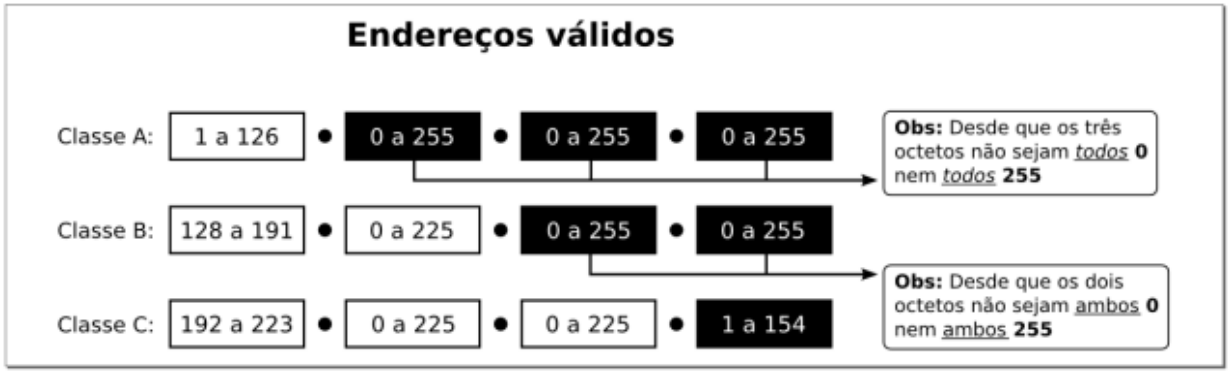
Continuando, temos a configuração das máscaras de sub-rede, que servem para indicar em que ponto termina a identificação da rede e começa a identificação do host. Ao usar a máscara "255.255.255.0", por exemplo, indicamos que os três primeiros números (ou octetos) do endereço servem para identificar a rede e apenas o último indica o endereço do host dentro dela.

Como vimos, na divisão original (que não é mais usada hoje em dia, como veremos a seguir) os endereços das três faixas eram diferenciados pelo número usado no primeiro octeto. Os endereços de classe A começavam com números de 1 a 126 (como, por exemplo, "62.34.32.1"), com máscara 255.0.0.0. Cada faixa de endereços classe A era composta de mais de 16 milhões de endereços mas, como existiam apenas 126 delas, elas eram reservadas para o uso de grandes empresas e órgãos governamentais.

Em seguida tínhamos os endereços de classe B, que englobavam os endereços iniciados com de 128 a 191, com máscara 255.255.0.0 (criando faixas compostas por 65 mil endereços) e o "terceiro mundo", que eram as faixas de endereços classe C. Elas abrangiam os endereços que começam com números de 192 a 223. As faixas de endereços de classe C eram mais numerosas, pois utilizavam máscara 255.255.255.0, mas, em compensação, cada faixa de classe C era composta por apenas 254 endereços. Veja alguns exemplos:

Ex. de endereço IP	Classe do endereço	Parte referente à rede	Parte referente ao host	Máscara de sub-rede padrão
98.158.201.128	Classe A	98.	158.201.128	255.0.0.0 (rede.host.host.host)
158.208.189.45	Classe B	158.208.	189.45	255.255.0.0 (rede.rede.host.host)
208.183.34.89	Classe C	208.183.34.	89	255.255.255.0 (rede.rede.rede.host)

Ao alugar um backbone vinculado a uma faixa de endereços classe C, por exemplo, você receberia uma faixa de endereços como "203.107.171.x", onde o "203.107.171" é o endereço de sua rede dentro da Internet, e o "x" é a faixa de 254 endereços que você pode usar para identificar seus servidores e os hosts dentro da rede. Na ilustração temos um resumo das regras para endereços TCP/IP válidos:



Como você pode notar no diagrama, nem todas as combinações de endereços são permitidas, pois o primeiro endereço (0) é reservado à identificação da rede, enquanto o último (255) é reservado ao endereço de broadcast, que é usado quando alguma estação precisa enviar um pacote simultaneamente para todos os micros dentro do segmento de rede.

Os pacotes de broadcast são usados para, por exemplo, configurar a rede via DHCP e localizar os compartilhamentos de arquivos dentro de uma rede Windows (usando o antigo protocolo NetBIOS). Mesmo os switches e hub-switches detectam os pacotes de broadcast e os transmitem simultaneamente para todas as portas. A desvantagem é que, se usados extensivamente, eles prejudicam o desempenho da rede.

Veja alguns exemplos de endereços **inválidos**:

0.xxx.xxx.xxx: Nenhum endereço IP pode começar com zero, pois ele é usado para o endereço da rede. A única situação em que um endereço começado com zero é usado, é quando um servidor DHCP responde à requisição da estação. Como ela ainda não possui um endereço definido, o pacote do servidor é endereçado ao endereço MAC da estação e ao endereço IP "0.0.0.0", o que faz com que o switch o envie para todos os micros da rede.

127.xxx.xxx.xxx: Nenhum endereço IP pode começar com o número 127, pois essa faixa de endereços é reservada para testes e para a interface de loopback. Se por exemplo você tiver um servidor de SMTP e configurar seu programa de e-mail para usar o servidor 127.0.0.1, ele acabará usando o servidor instalado na sua própria máquina. O mesmo acontece ao tentar acessar o endereço 127.0.0.1 no navegador: você vai cair em um servidor web habilitado na sua máquina. Além de testes em geral, a interface de loopback é usada para comunicação entre diversos programas, sobretudo no Linux e outros sistemas Unix.

255.xxx.xxx.xxx, xxx.255.255.255, xxx.xxx.255.255: Nenhum identificador de rede pode ser 255 e nenhum identificador de host pode ser composto apenas de endereços 255, seja qual for a classe do endereço, pois estes endereços são usados para enviar pacotes de

broadcast. Outras combinações são permitidas, como em 65.34.255.197 (em um endereço de classe A) ou em 165.32.255.78 (endereço de classe B).

xxx.0.0.0, xxx.xxx.0.0: Nenhum identificador de host pode ser composto apenas de zeros, seja qual for a classe do endereço, pois estes endereços são reservados para o endereço da rede. Como no exemplo anterior, são permitidas outras combinações como 69.89.0.129 (classe A) ou 149.34.0.95 (classe B).

xxx.xxx.xxx.255, xxx.xxx.xxx.0: Nenhum endereço de classe C pode terminar com 0 ou com 255, pois, como já vimos, um host não pode ser representado apenas por valores 0 ou 255, já que eles são usados para o envio de pacotes de broadcast.

Dentro de redes locais, é possível usar máscaras diferentes para utilizar os endereços IP disponíveis de formas diferentes das padrão. O importante neste caso é que todos os micros da rede sejam configurados com a mesma máscara, caso contrário você terá problemas de conectividade, já que tecnicamente os micros estarão em redes diferentes.

Um exemplo comum é o uso da faixa de endereços 192.168.0.x para redes locais. Originalmente, esta é uma faixa de endereços classe C e por isso a máscara padrão é 255.255.255.0. Mesmo assim, muita gente prefere usar a máscara 255.255.0.0, o que permite mudar os dois últimos octetos (192.168.x.x). Neste caso, você poderia ter dois micros, um com o IP "192.168.2.45" e o outro com o IP "192.168.34.65" e ambos se enxergariam perfeitamente, pois entenderiam que fazem parte da mesma rede. Não existe problema em fazer isso, desde que você use a mesma máscara em todos os micros da rede.

A divisão tradicional, com as classes A, B e C de endereços IP fazia com que um grande número de endereços fossem desperdiçados. Um provedor de acesso que precisasse de 10.000 endereços IP, por exemplo, precisaria ou utilizar uma faixa de endereços classe B inteira (65 mil endereços), o que geraria um grande desperdício, ou utilizar 40 faixas de endereços classe C separadas, o que complicaria a configuração. Existia ainda o problema com as faixas de endereços classe A, que geravam um brutal desperdício de endereços, já que nenhuma empresa ou organização sozinha chega a utilizar 16 milhões de endereços IP.

A solução para o problema foi a implantação do sistema CIDR (abreviação de "Classless Inter-Domain Routing", que pronunciamos como "cider"), a partir de 1993 (leia o RCF no http://tools.ietf.org/html/rfc1519).

Entender as classes de endereços A, B e C é importante para compreender o uso das máscaras de sub-rede e por isso elas ainda são muito estudadas, mas é importante ter em mente que, na prática, elas são uma designação obsoleta. Naturalmente, ainda existem muitas redes que utilizam faixas de endereços de classe A, B e C (já que as faixas alocadas no passado não podem ser simplesmente revogadas de uma hora para a outra), mas as faixas alocadas atualmente utilizam quase sempre o novo sistema.

No CIDR são utilizadas máscaras de tamanho variável (o termo em inglês é VLSM, ou Variable-Length Subnet Mask), que permitem uma flexibilidade muito maior na criação das faixas de endereços. Se são necessários apenas 1000 endereços, por exemplo, poderia ser usada uma máscara /22 (que permite o uso de 1022 endereços), em vez de uma faixa de classe B inteira, como seria necessário antigamente.

Outra mudança é que as faixas de endereços não precisam mais iniciar com determinados números. Uma faixa com máscara /24 (equivalente a uma faixa de endereços de classe C) pode começar com qualquer dígito e não apenas com de 192 a 223.

O CIDR permite também que várias faixas de endereços contínuas sejam agrupadas em faixas maiores, de forma a simplificar a configuração. É possível agrupar 8 faixas de endereços com máscara 255.255.255.0 (classe C) contínuas em uma única faixa com máscara /21, por exemplo, que oferece um total de 2045 endereços utilizáveis (descontando o endereço da rede, endereço de broadcast e o endereço do gateway).

As faixas de endereços são originalmente atribuídas pela IANA às entidades regionais. Elas dividem os endereços em faixas menores e as atribuem aos carriers (as operadoras responsáveis pelos links), empresas de hospedagem, provedores de acesso e outras instituições. Estas, por sua vez, quebram os endereços em faixas ainda menores, que são atribuídas aos consumidores finais.

Revisando, a máscara de subrede determina qual parte do endereço IP é usada para endereçar a rede e qual é usada para endereçar os hosts dentro dela. No endereço 200.232.211.54, com máscara 255.255.255.0 (/24), por exemplo, os primeiros 24 bits (200.232.211.) endereçam a rede e os 8 últimos (54) endereçam o host.

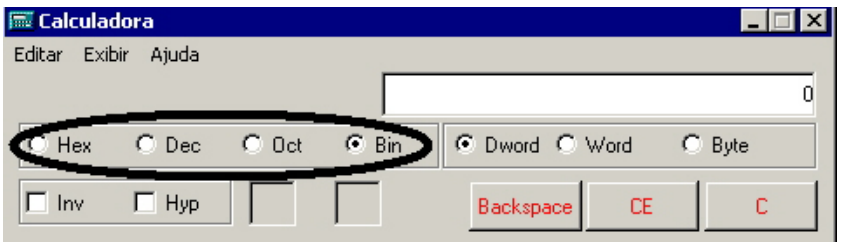
Quando usamos máscaras simples, podemos trabalhar com os endereços em decimais, pois são sempre reservados 1, 2 ou 3 octetos inteiros para a rede e o que sobra fica reservado ao host. Esta é a idéia usada nas faixas de endereços classe A, B e C.

Quando falamos em máscaras de tamanho variável, entretanto, precisamos começar a trabalhar com endereços binários, pois a divisão pode ser feita em qualquer ponto. Imagine, por exemplo, o endereço "72.232.35.108". Originalmente, ele seria um endereço de classe A e utilizaria máscara "255.0.0.0". Mas, utilizando máscaras de tamanho variável, ele poderia utilizar a máscara "255.255.255.248", por exemplo.

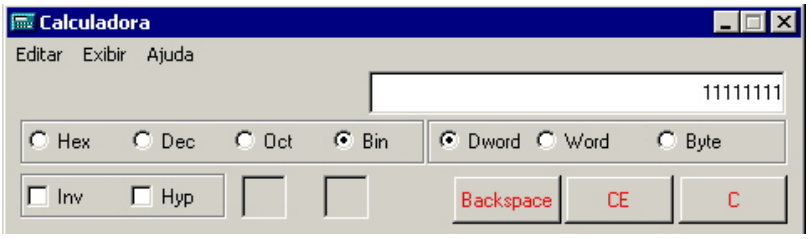
Nesse caso, teríamos 29 bits do endereço dedicados à endereçar a rede e apenas os 3 últimos bits destinados ao host. Convertendo o endereço para binário teríamos o endereço "01001000.11101000.01100000.01101100", onde o "01001000.11101000.01100000.01101" é o endereço da rede e o "100" é o endereço do host dentro dela. Como temos 29 bits dedicados à rede, é comum o uso de um "/29" como máscara, no lugar de "255.255.255.248".

À primeira vista, esse conceito parece bastante complicado, mas na prática não é tão difícil assim. A primeira coisa a ter em mente é que as máscaras de tamanho variável só fazem sentido quando você converte o endereço IP para binário.

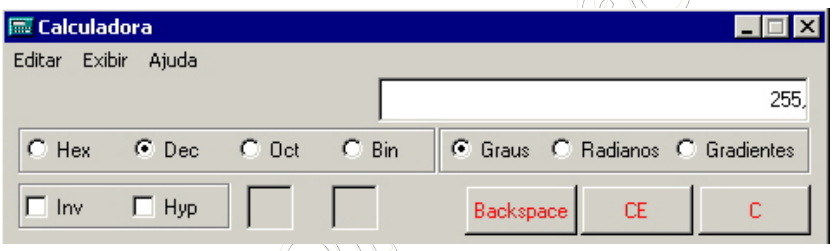
Para converter um número decimal em um número binário, você pode usar a calculadora do Windows ou o Kcalc no Linux. Configure a calculadora para o modo científico (exibir/científica) e verá que do lado esquerdo aparecerá um menu de seleção permitindo (entre outras opções) escolher entre decimal (dec) e binário (bin).



Configure a calculadora para binário e digite o número 11111111, mude a opção da calculadora para decimal (dec) e a calculadora mostrará o número 255, que é o seu correspondente em decimal. Tente de novo agora com o binário 00000000 e terá o número decimal 0.



Veja que 0 e 255 são exatamente os números que usamos nas máscaras de sub-rede simples.



O número decimal 255 (equivalente a 11111111) indica que todos os 8 números binários do octeto se referem à rede, enquanto o decimal 0 (correspondente a 00000000) indica que todos os 8 dígitos binários do octeto se referem ao host. Em uma rede com máscara 255.255.255.0 temos:

Decimal:	**255**	**255**	**255**	**0**
Binário:	**11111111**	**11111111**	**11111111**	**00000000**
	rede	**rede**	**rede**	**host**

As máscaras de tamanho variável permitem fazer a divisão em outros pontos do endereço. No endereço "72.232.35.108" com máscara "255.255.255.248" que citei a pouco, teríamos:

Decimal:	**255**	**255**	**255**	**248**
Binário:	**11111111**	**11111111**	**11111111**	**11111 000**
	rede	**rede**	**rede**	**rede host**

Imagine o caso de um pequeno provedor de acesso, que possui um backbone com uma faixa de endereços com máscara 255.255.255.0 (/24) e precisa dividí-lo entre dois clientes, onde cada um deles deve ter uma faixa completa de endereços.

O backbone do provedor utiliza a faixa de endereços 203.107.171.x onde o 203.107.171 é o endereço da rede e o "x" é a faixa de endereços de que eles dispõem para endereçar os micros das duas empresas. Como endereçar ambas as redes, se não é possível alterar o "203.107.171" que é a parte do seu endereço que se refere à rede? A solução seria justamente utilizar máscaras de tamanho variável.

Veja que podemos alterar apenas dos últimos 8 bits do endereço IP:

Decimal:	**203**	**107**	**171**	**x**
Binário:	**11001011**	**11010110**	**10101011**	**????????**

Usando uma máscara 255.255.255.0, são reservados todos os 8 bits para o endereçamento dos hosts, e não sobra nada para diferenciar as duas redes. Usando uma máscara de tamanho variável, é possível "quebrar" os 8 bits do octeto em duas partes, usando a primeira para diferenciar as duas redes e a segunda para endereçar os hosts:

Decimal:	**203**	**107**	**171**	**x**
Binário:	**11001011**	**11010110**	**10101011**	**???? ????**
	rede	**rede**	**rede**	**rede host**

Para tanto, ao invés de usar a máscara de sub-rede 255.255.255.0 que, como vimos, reservaria todos os 8 bits para o endereçamento do host, usaremos uma máscara 255.255.255.240 (/28) (corresponde ao binário 11111111.111111.11111111.11110000). Veja que em uma máscara de sub-rede os números binários "1" referem-se à rede e os números "0" referem-se ao host. Na máscara 255.255.255.240 temos exatamente esta divisão: quatro bits 1 e quatro bits 0:

Decimal:	**255**	**255**	**255**	**240**
Binário:	**11111111**	**11111111**	**11111111**	**1111 0000**
	rede	**rede**	**rede**	**rede host**

Temos agora o último octeto dividido em dois endereços binários de 4 bits cada. Cada um dos dois grupos representa agora um endereço distinto, e deve ser configurado independentemente. Como fazer isso? Veja que 4 bits permitem 16 combinações diferentes (de 0 a 15). Se você converter o número 15 em binário terá "1111" e, se converter o decimal 0, terá "0000". Se converter o decimal 11 terá "1011" e assim por diante.

Neste caso, é possível usar endereços de 1 a 14 para identificar os hosts e as redes separadas. Note que os endereços 0 e 15 não podem ser usados, pois assim como os endereços 0 e 255, eles são reservados para pacotes de broadcast:

Decimal:	**203**	**107**	**171**	**12 _ 14**
Binário:	**11111111**	**11111111**	**11111111**	**1100 1110**
	rede	**rede**	**rede**	**rede host**

Estabeleça um endereço de rede para cada uma das duas sub-redes disponíveis e um endereço diferente para cada micro da rede, mantendo a formatação do exemplo anterior. Por enquanto, apenas anote em um papel os endereços escolhidos, junto como seu correspondente em binários.

Na hora de configurar o endereço IP nas estações, configure primeiro a máscara de sub-rede como 255.255.255.240 e, em seguida, converta os endereços binários em decimais, para ter o endereço IP de cada estação. No exemplo da ilustração anterior, havíamos estabelecido o endereço 12 para a rede e o endereço 14 para a estação; 12 corresponde a "1100" e 14 corresponde a "1110". Juntando os dois temos "11001110", que corresponde ao decimal "206". O endereço IP da estação será então 203.107.171.206, com máscara 255.255.255.240.

Se tivesse escolhido o endereço 10 para a rede e o endereço 8 para a estação, teríamos "10101000" que corresponde ao decimal 168. Neste caso, o endereço IP da estação seria 203.107.171.168.

Nesse primeiro exemplo dividimos a faixa de endereços em 14 redes distintas, cada uma com 14 endereços. Isso permitiria que o provedor de acesso do exemplo fornecesse links para até 14 empresas diferentes, desde que cada uma não precisasse de mais do que 14 endereços. É possível criar diferentes combinações, reservando números diferentes de bits para a rede e o host:

Máscara	Bits da rede	Bits do host	Número de redes	Número de hosts
255.255.255.0 (/24)	nenhum	00000000	nenhuma	254 endereços (do 1 ao 254)
255.255.255.192 (/26)	11	000000	2 endereços (2 e 3)	62 endereços (de 1 a 62)
255.255.255.224 (/27)	111	00000	6 endereços (de 1 a 6)	30 endereços (de 1 a 30)
255.255.255.240 (/28)	1111	0000	14 endereços (de 1 a 14)	14 endereços (de 1 a 14)
255.255.255.248 (/29)	11111	000	30 endereços (de 1 a 30)	6 endereços (de 1 a 6)
255.255.255.252 (/30)	111111	00	62 endereços (de 1 a 62)	2 endereços (2 e 3)

Em qualquer um dos casos, para obter o endereço IP basta converter os dois endereços (rede e estação) para binário, "juntar" os bits e converter o octeto para decimal.

Usando uma máscara de sub-rede 192, por exemplo, e estabelecendo o endereço 2 (ou "10" em binário) para a rede e 47 (ou "101111" em binário) para o host, juntaríamos ambos os binários obtendo o octeto "10101111" que corresponde ao decimal "175".

Se usássemos a máscara de sub-rede 248, estabelecendo o endereço 17 (binário "10001") para a rede e o endereço 5 (binário "101") para o host, obteríamos o octeto "10001101" que corresponde ao decimal "141".

Na hora de escrever o endereço e a máscara (como ao criar uma regra de firewall, ou ajustar um determinado arquivo com permissões de acesso), você pode tanto escrever a máscara por extenso, como em "192.168.0.0/255.255.255.0", quanto escrever usado a notação abreviada, como em "192.168.0.0/24".

Essa possibilidade é usada ao extremo pelas empresas de hospedagem, que dividem as faixas de endereços disponíveis em diversas faixas menores, com apenas 4 ou 8 endereços, que são atribuídas aos servidores dedicados hospedados em seus data-centers.

Ao usar a máscara 255.255.255.248, por exemplo, apenas 3 bits do endereço são reservados ao endereçamento dos hosts (convertendo 255.255.255.248 para binário, você teria

11111111.11111111.11111111.11111000), permitindo que uma faixa de endereços originalmente de classe A (16 milhões de hosts) seja dividida em 2.080.768 pequenas redes, uma para cada servidor dedicado que for locado.

Três bits permitem 8 combinações, mas o primeiro e o último endereço são reservados ao endereço da rede e ao endereço de broadcast, fazendo com que apenas 6 endereços possam realmente ser utilizados. Destes, mais um é sacrificado, pois é atribuído ao gateway (sem o gateway o servidor não acessa a Internet), de forma que no final apenas 5 endereços ficam realmente disponíveis.

Imagine, por exemplo, que você locou um servidor dedicado que tem disponível uma faixa de endereços que vai do 72.232.35.106 até o 72.232.35.110 (com máscara 255.255.255.248), sendo que o endereço 72.232.35.105 é o gateway da rede. Originalmente, a faixa de endereços iria do 72.232.35.104 ao 72.232.35.111, mas como o primeiro endereço é reservado à rede, o último aos pacotes de broadcast e mais um endereço precisa ficar reservado ao gateway da rede, ficamos no final com apenas 5 endereços válidos, como citei. Convertendo os endereços para binário, teríamos:

```
72.232.35.104 = 01001000 . 11101000 . 00100011 . 01101 000
72.232.35.105 = 01001000 . 11101000 . 00100011 . 01101 001
72.232.35.106 = 01001000 . 11101000 . 00100011 . 01101 010
72.232.35.107 = 01001000 . 11101000 . 00100011 . 01101 011
72.232.35.108 = 01001000 . 11101000 . 00100011 . 01101 100
72.232.35.109 = 01001000 . 11101000 . 00100011 . 01101 101
72.232.35.110 = 01001000 . 11101000 . 00100011 . 01101 110
72.232.35.111 = 01001000 . 11101000 . 00100011 . 01101 111
```

Como pode ver, os 8 endereços esgotam todas as possibilidades possíveis dentro da faixa, já que temos apenas 3 bits disponíveis para o host. Os 29 primeiros bits do endereço se referem à rede e por isso são sempre iguais e apenas os três últimos se referem ao host. Este processo de converter os endereços para binário é um pouco trabalhoso, mas ajuda a entender melhor a estrutura dos endereços no CIDR.

Você pode se perguntar qual é a necessidade de ter uma faixa com 5 endereços utilizáveis se o servidor é apenas um. Existem diversos motivos para isso. A primeira é que, ao configurar um servidor dedicado, você precisa de uma faixa de endereços inteira para poder configurar o DNS reverso, um pré-requisito para que seus e-mails não sejam rotulados como spam por outros servidores.

Ao registrar um domínio, você precisa fornecer os endereços de dois servidores DNS, que responderão por ele. Em vez de ter dois servidores, você pode utilizar outro dos seus 5 endereços disponíveis para criar um alias (apelido) para a placa de rede do seu servidor dedicado e assim poder configurá-lo para responder simultaneamente como servidor DNS primário e secundário, eliminando assim a necessidade de utilizar dois servidores separados. Novamente, essa configuração é possível apenas caso o servidor possua uma faixa de endereços própria.

No final, a configuração de rede de um servidor dedicado acaba sendo algo similar a isto:

```
Endereço IP: 72.232.35.106
Máscara: 255.255.255.248
Gateway: 72.232.35.105
Endereço da rede: 72.232.35.104
Endereço de broadcast: 72.232.35.111
Alias da placa de rede (para o DNS secundário): 72.232.35.107
Endereços vagos: 72.232.35.108, 72.232.35.109 e 72.232.35.110
```

9.5.2 Cálculo IP

Vou iniciar falando do sistema de numeração decimal, para depois fazer uma analogia ao apresentar o sistema de numeração binário.Todos nos conhecemos o sistema de numeração decimal, no qual são baseados os números que usamos no nosso dia-a-dia, como por exemplo: 100, 259, 1450 e assim por diante. Você já parou para pensar porque este sistema de numeração é chamado de sistema de numeração decimal? Não? Bem, a resposta é bastante simples: este sistema é baseado em dez dígitos diferentes, por isso é chamado de sistema de numeração decimal. Todos os números do sistema de numeração decimal são escritos usando-se uma combinação dos seguintes dez dígitos:

0 1 2 3 4 5 6 7 8 9

Dez dígitos **-> Sistema de numeração decimal.**

Vamos analisar como é determinado o valor de um número do sistema de numeração decimal. Por exemplo, considere o seguinte número: **4538**

O valor deste número é formado, multiplicando-se os dígitos do número, de trás para frente, por potências de 10, começando com 10^0. O último dígito (bem à direita) é multiplicado por 10^0, o penúltimo por 10^1, o próximo por 10^2 e assim por diante. O valor real do número é a soma dos resultados destas multiplicações. Observe o esquema a seguir que será bem mais fácil de entender:

	4	**5**	**3**	**8**
Multiplica por:	**10^3**	**10^2**	**10^1**	**10^0**
ou seja:	1000	100	10	1
Resultado:	4x1000	5x100	3x10	8x1
Igual a:	4000	500	30	8
Somando tudo:	**4000+500+30+8**			
É igual a:	**4538**			

Observe que 4538 significa exatamente:

4 milhares (10^3)
+ 5 centenas (10^2)
+ 3 dezenas (10^1)
+ 8 unidades (10^0)

E assim para números maiores, com mais dígitos, teríamos potências de 10^4, 10^5 e assim por diante. Observe que multiplicando cada dígito por potências de 10, obtemos o número original. Este princípio aplicado ao sistema de numeração decimal é válido para qualquer sistema de numeração. Se for o sistema de numeração Octal (baseado em 8 dígitos), multiplica-se por potências de 8: 8^0, 8^1, 8^2 e assim por diante. Se for o sistema Hexadecimal (baseado em 10 dígitos e 6 letras) multiplica-se por potências de 16, só que a letra A equivale a 10, já que não tem sentido multiplicar por uma letra, a letra B equivale a 11 e assim por diante.

Bem, por analogia, se o sistema decimal é baseado em dez dígitos, então o sistema binário deve ser baseado em dois dígitos? Exatamente. Os números no sistema binários são escritos usando-se apenas os dois seguintes dígitos: 0 1

Isso mesmo, números no sistema binário são escritos usando-se apenas zeros e uns, como nos exemplos a seguir: **01011100 , 11011110 , 00011111**

Também por analogia, se, no sistema decimal, para obter o valor do número, multiplicamos os seus dígitos, de trás para frente, por potências de 10, no sistema binário fizemos esta mesma operação, só que baseada em potências de 2, ou seja: 2^0, 2^1, 2^2, 2^3, 2^4 e assim por diante.

Vamos considerar alguns exemplos práticos. Como faço para saber o valor decimal do seguinte número binário: **11001110**

Vamos utilizar a tabelinha a seguir para facilitar os nossos cálculos:

	1	**1**	**0**	**0**	**1**	**1**	**1**	**0**
Multiplica por:	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
equivale a:	128	64	32	16	8	4	2	1
Multiplicação:	1**x**128	1**x**64	0**x**32	0**x**16	1**x**8	1**x**4	1**x**2	0**x**1
Resulta em:	128	64	0	0	8	4	2	0
Somando tudo:	**128+64+0+0+8+4+2+0**							
Resulta em:	**206**							

Ou seja, o número binário 11001110 equivale ao decimal 206. Observe que onde temos um a respectiva potência de 2 é somada e onde temos o zero a respectiva potência de 2 é anulada por ser multiplicada por zero. Apenas para fixar um pouco mais este conceito, vamos fazer mais um exemplo de conversão de binário para decimal. Converter o número **11100010** para decimal:

	1	**1**	**1**	**0**	**0**	**0**	**1**	**0**
Multiplica por:	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
equivale a:	128	64	32	16	8	4	2	1
Multiplicação:	1**x**128	1**x**64	1**x**32	0**x**16	0**x**8	0**x**4	1**x**2	0**x**1
Resulta em:	128	64	32	0	0	0	2	0
Somando tudo:	**128+64+32+0+0+0+2+0**							
Resulta em:	**226**							

Bem, e se tivéssemos que fazer o contrário, converter o número 234 de decimal para binário, qual seria o binário equivalente??

Nota: Nos exemplos vou trabalhar com valores decimais de, no máximo, 255, que são valores que podem ser representados por 8 dígitos binários, ou na linguagem do computador 8 bits, o que equivale exatamente a um byte. Por isso que cada um dos quatro números que fazem parte do número IP, somente podem ter um valor máximo de 255, que é um valor que cabe em um byte, ou seja, 8 bits.

Existem muitas regras para fazer esta conversão, eu prefiro utilizar uma bem simples, que descreverei a seguir e que serve perfeitamente para o propósito deste tutorial.

Vamos voltar ao nosso exemplo, como converter 234 para um binário de 8 dígitos?

Eu começo o raciocínio assim. Primeiro vamos lembrar o valor decimal correspondente a cada um dos oito dígitos binários: 128 64 32 16 8 4 2 1

Lembrando que estes números representam potências de 2, começando, de trás para frente, com 2^0, 2^1, 2^2 e assim por diante, conforme indicado logo a seguir:

128 64 32 16 8 4 2 1
2^7 2^6 2^5 2^4 2^3 2^2 2^1 2^0

Pergunto: 128 cabe em 234? Sim, então o primeiro dígito é 1. Somando 64 a 128 passa de 234? Não, dá 192, então o segundo dígito também é 1. Somando 32 a 192 passa de 234? Não, dá 224, então o terceiro dígito também é 1. Somando 16 a 224 passa de 234? Passa, então o quarto dígito é zero. Somando 8 a 224 passa de 234? Não, da 232, então o quinto dígito é 1. Somando 4 a 232 passa de 234? Passa, então o sexto dígito é zero. Somando 2 a 232 passa de 234? Não, dá exatamente 234, então o sétimo dígito é 1. Já cheguei ao valor desejado, então todos os demais dígitos são zero. Com isso, o valor 234 em binário é igual a: 11101010

Para exercitar vamos converter mais um número de decimal para binário. Vamos converter o número 144 para binário.

Pergunto: 128 cabe em 144? Sim, então o primeiro dígito é 1. Somando 64 a 128 passa de 144? Sim, dá 192, então o segundo dígito é 0. Somando 32 a 128 passa de 144? Sim, dá 160, então o terceiro dígito também é 0. Somando 16 a 128 passa de 144? Não, dá exatamente 144, então o quarto dígito é 1. Já cheguei ao valor desejado, então todos os demais dígitos são zero. Com isso, o valor 144 em binário é igual a: **10010000**

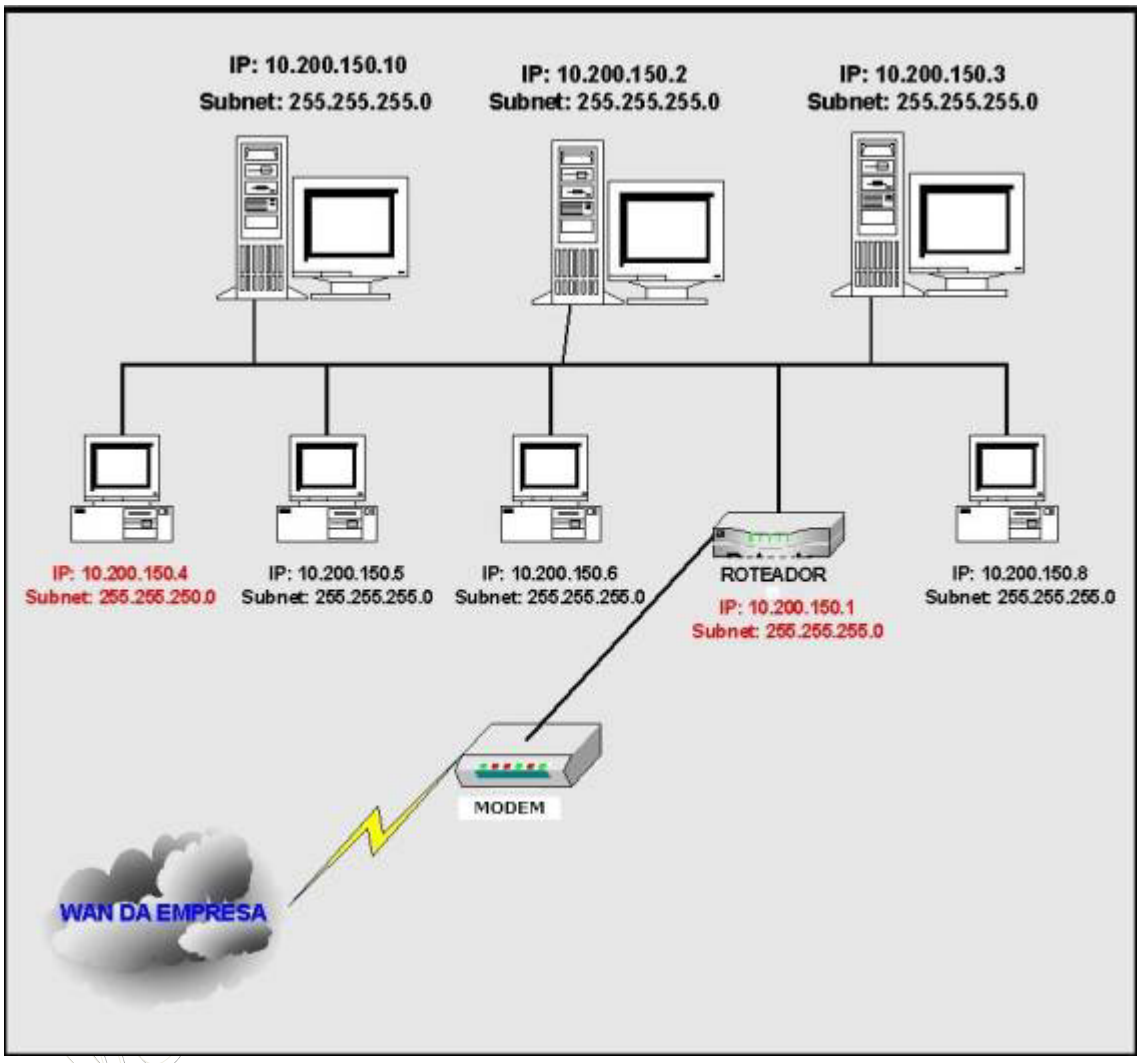
Bem, agora que você já sabe como converter de decimal para binário, está em condições de aprender sobre o operador "**E**" e como o TCP/IP usa a máscara de sub-rede (subnet mask) e uma operação "E", para verificar se duas máquinas estão na mesma rede ou em redes diferentes.

Existem diversas operações lógicas que podem ser feitas entre dois dígitos binários, sendo as mais conhecidas as seguintes: "E", "OU", "XOR" e "NOT".

Para o nosso estudo interessa o operador E. Quando realizamos um "E" entre dois bits, o resultado somente será 1, se os dois bits forem iguais a 1. Se pelo menos um dos bits for igual a zero, o resultado será zero. Na tabela a seguir temos todos os valores possíveis da operação E entre dois bits:

bit-1	bit-2	(bit-1) E (bit-2)
1	**1**	**1**
1	**0**	**0**
0	**1**	**0**
0	**0**	**0**

Vamos exemplificar como o TCP/IP usa a máscara de rede. Considere a figura a seguir, onde temos a representação de uma rede local, ligada a outras redes da empresa, através de um roteador.



Temos uma rede que usa como máscara de sub-rede 255.255.255.0. A rede é a 10.200.150.0, ou seja, todos os equipamentos da rede tem os três primeiras partes do número IP como sendo: 10.200.150. Veja que existe uma relação direta entre a máscara de sub-rede a quantas das partes do número IP são fixas, ou seja, que definem a rede, conforme foi descrito na Parte 1 deste curso.

A rede da figura anterior é uma rede das mais comumente encontradas hoje em dia, onde existe um roteador ligado à rede e o roteador está conectado a um Modem, através do qual é feita a conexão da rede local com a rede WAN da empresa, através de uma linha de dados (também conhecido como link de comunicação). Nas próximas partes lições vou detalhar a função do roteador e mostrarei como funciona o roteamento entre redes.

Quando dois computadores tentam trocar informações em uma rede, o TCP/IP precisa, primeiro, determinar se os dois computadores pertencem a mesma rede ou a redes diferentes. Neste caso podemos ter duas situações distintas:

Situação 1: Os dois computadores pertencem a mesma rede: Neste caso o TCP/IP envia o pacote para o barramento local da rede. Todos os computadores recebem o pacote, mas somente o computador que é o destinatário do pacote é que o captura e passa para processamento pelo Windows e

pelo programa de destino. Como é que o computador sabe se ele é ou não o destinatário do pacote? Muito simples, no pacote de informações está contido o endereço IP do computador destinatário. Em cada computador, o TCP/IP compara o IP de destinatário do pacote com o IP do computador, para saber se o pacote é ou não para o respectivo computador.

Situação 2: Os dois computadores não pertencem a mesma rede: Neste caso o TCP/IP envia o pacote para o Roteador (endereço do Default Gateway configurado nas propriedades do TCP/IP) e o Roteador se encarrega de fazer o pacote chegar ao seu destino. Em uma das partes deste tutorial veremos detalhes sobre como o Roteador é capaz de rotear pacotes de informações até redes distantes.

Agora a pergunta que tem a ver com este tópico:

"Como é que o TCP/IP faz para saber se o computador de origem e o computador de destino pertencem a mesma rede?"

Vamos usar alguns exemplos práticos para explicar como o TCP/IP faz isso:

Exemplo 1: Com base na figura anterior, suponha que o computador cujo IP é 10.200.150.5 (origem) queira enviar um pacote de informações para o computador cujo IP é 10.200.150.8 (destino), ambos com máscara de sub-rede igual a 255.255.255.0.

O primeiro passo é converter o número IP das duas máquinas e da máscara de sub-rede para binário. Com base nas regras que vimos anteriormente, teríamos a seguinte conversão:

Computador de origem:

10	**200**	**150**	**5**
00001010	11001000	10010110	00000101

Computador de destino:

10	**200**	**150**	**8**
00001010	11001000	10010110	00001000

Máscara de sub-rede:

255	**255**	**255**	**0**
11111111	11111111	11111111	00000000

Feitas as conversões para binário, vamos ver que tipo de cálculos o TCP/IP faz, para determinar se o computador de origem e o computador de destino estão na mesma rede.

Em primeiro lugar é feita uma operação "E", bit a bit, entre o Número IP e a máscara de Sub-rede do computador de origem, conforme indicado na tabela a seguir:

10.200.150.5	00001010	11001000	10010110	00000101	**E**
255.255.255.0	11111111	11111111	11111111	00000000	
10.200.150.0	00001010	11001000	10010110	00000000	**Resultado**

Agora é feita uma operação "E", bit a bit, entre o Número IP e a máscara de sub-rede do computador de destino, conforme indicado na tabela a seguir:

10.200.150.8	00001010	11001000	10010110	00001000	**E**
255.255.255.0	11111111	11111111	11111111	00000000	
10.200.150.0	00001010	11001000	10010110	00000000	**Resultado**

Agora o TCP/IP compara os resultados das duas operações. Se os dois resultados forem iguais, aos dois computadores, origem e destino, pertencem a mesma rede local. Neste caso o TCP/IP envia o pacote

para o barramento da rede local. Todos os computadores recebem o pacote, mas somente o destinatário do pacote é que o captura e passa para processamento pelo Windows e pelo programa de destino. Como é que o computador sabe se ele é ou não o destinatário do pacote? Muito simples, no pacote de informações está contido o endereço IP do destinatário. Em cada computador, o TCP/IP compara o IP de destinatário do pacote com o IP do computador, para saber se o pacote é ou não para o respectivo computador.

É o que acontece neste exemplo, pois o resultado das duas operações "E" é igual: 10.200.150.0, ou seja, os dois computadores pertencem a rede: 10.200.150.0

Como você já deve ter adivinhado, agora vamos a um exemplo, onde os dois computadores não pertencem a mesma rede, pelo menos devido às configurações do TCP/IP.

Exemplo 2: Suponha que o computador cujo IP é 10.200.150.5 (origem) queira enviar um pacote de informações para o computador cujo IP é 10.204.150.8 (destino), ambos com máscara de sub-rede igual a 255.255.255.0.

O primeiro passo é converter o número IP das duas máquinas e da máscara de sub-rede para binário. Com base nas regras que vimos anteriormente, teríamos a seguinte conversão:

Computador de origem:

10	**200**	**150**	**5**
00001010	11001000	10010110	00000101

Computador de destino:

10	**204**	**150**	**8**
00001010	11001100	10010110	00001000

Máscara de sub-rede:

255	**255**	**255**	**0**
11111111	11111111	11111111	00000000

Feitas as conversões para binário, vamos ver que tipo de cálculos o TCP/IP faz, para determinar se o computador de origem e o computador de destino estão na mesma rede. Em primeiro lugar é feita uma operação "E", bit a bit, entre o Número IP e a máscara de Sub-rede do computador de origem, conforme indicado na tabela a seguir:

10.200.150.5	00001010	11001000	10010110	00000101	**E**
255.255.255.0	11111111	11111111	11111111	00000000	
10.200.150.0	00001010	11001000	10010110	00000000	**Resultado**

Agora é feita uma operação "E", bit a bit, entre o Número IP e a máscara de sub-rede do computador de destino, conforme indicado na tabela a seguir:

10.204.150.8	00001010	11001100	10010110	00001000	**E**
255.255.255.0	11111111	11111111	11111111	00000000	
10.204.150.0	00001010	11001100	10010110	00000000	**Resultado**

Agora o TCP/IP compara os resultados das duas operações. Neste exemplo, os dois resultados são diferentes: 10.200.150.0 e 10.204.150.0. Nesta situação o TCP/IP envia o pacote para o Roteador (endereço do Default Gateway configurado nas propriedades do TCP/IP) e o Roteador se encarrega de fazer o pacote chegar a rede do computador de destino. Em outras palavras o Roteador sabe entregar o pacote para a rede 10.204.150.0 ou sabe para quem enviar (um outro roteador), para que este próximo

roteador possa encaminhar o pacote. Este processo continua até que o pacote seja entregue na rede de destino ou seja descartado, por não ter sido encontrada uma rota para a rede de destino.

Observe que, na figura anterior, temos dois computadores que, apesar de estarem fisicamente na mesma rede, não conseguirão se comunicar devido a um erro de configuração na máscara de sub-rede de um dos computadores. É o caso do computador 10.200.150.4 (com máscara de sub-rede 255.255.250.0). Como este computador está com uma máscara de sub-rede diferente dos demais computadores da rede (255.255.255.0), ao fazer os cálculos, o TCP/IP chega a conclusão que este computador pertence a uma rede diferente, o que faz com que ele não consiga se comunicar com os demais computadores da rede local.

9.5.3 Roteamento

Vimos que a máscara de sub-rede é utilizada para determinar qual "parte" do endereço IP representa o número da Rede e qual parte representa o número da máquina dentro da rede. A máscara de sub-rede também foi utilizada na definição original das classes de endereço IP. Em cada classe existe um determinado número de redes possíveis e, em cada rede, um número máximo de máquinas. Com base na máscara de sub-rede o protocolo TCP/IP determina se o computador de origem e o de destino estão na mesma rede local. Com base em cálculos binários, o TCP/IP pode chegar a dois resultados distintos:

- **O computador de origem e o computador de destino estão na mesma rede local:** Neste caso os dados são enviados para o barramento da rede local. Todos os computadores da rede recebem os dados. Ao receber os dados cada computador analisa o campo Número IP do destinatário. Se o IP do destinatário for igual ao IP do computador, os dados são capturados e processados pelo sistema, caso contrário são simplesmente descartados. Observe que com este procedimento, apenas o computador de destino é que efetivamente processa os dados para ele enviados, os demais computadores simplesmente descartam os dados.

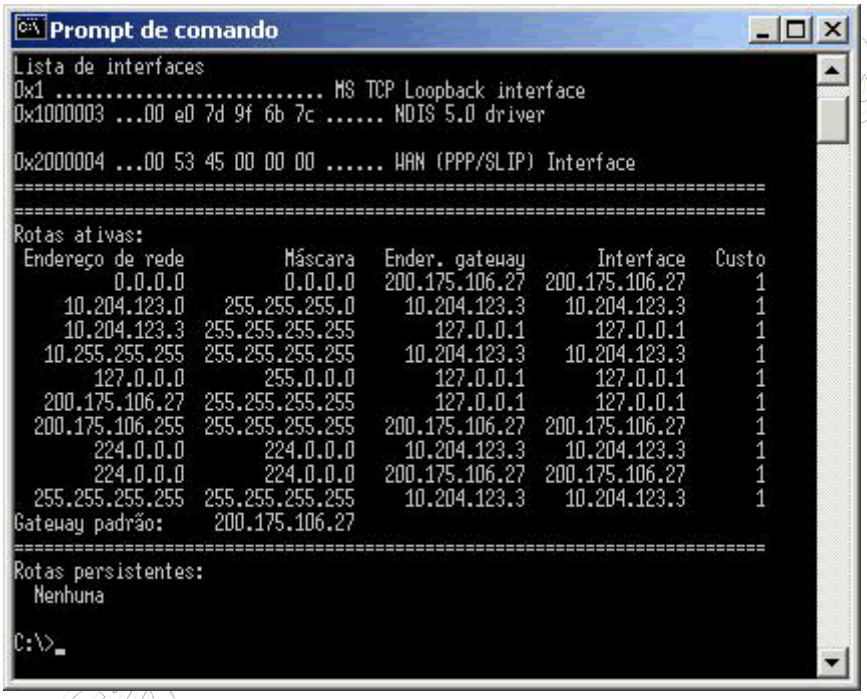
- **O computador de origem e de destino não estão na mesma rede local:** Neste caso os dados são enviados o equipamento com o número IP configurado no parâmetro Default Gateway (Gateway Padrão). Ou seja, se após os cálculos baseados na máscara de sub-rede, o TCP/IP chegar a conclusão que o computador de destino e o computador de origem não fazem parte da mesma rede local, os dados são enviados para o Default Gateway, o qual será encarregado de encontrar um caminho para enviar os dados até o computador de destino. Esse "encontrar o caminho" é tecnicamente conhecido como Rotear os dados até o destino (ou melhor, rotear os dados até a rede do computador de destino). O responsável por "Rotear" os dados é o equipamento que atua como Default Gateway o qual é conhecido como Roteador. Com isso fica fácil entender o papel do Roteador:

"O Roteador é o responsável por encontrar um caminho entre a rede onde está o computador que enviou os dados (computador de origem) e a rede onde está o computador que irá receber os dados (computador de destino)."

Quando ocorre um problema com o Roteador, tornando-o indisponível, você consegue se comunicar normalmente com os demais computadores da sua rede local, porém não conseguirá comunicação com outras redes de computadores, como por exemplo a Internet.

Toda a funcionalidade do Roteador é baseada em tabelas de roteamento. Quando um pacote chega em uma das interfaces do roteador, ele analisa a sua tabela de roteamento, para verificar se na tabela de roteamento, existe uma rota para a rede de destino. Pode ser uma rota direta ou então para qual roteador o pacote deve ser enviado. Este processo continua até que o pacote seja entregue na rede de destino, ou até que o limite de 16 hopes (para simplificar imagine um hope como sendo um roteador da rede) tenha sido atingido.

Na Figura a seguir apresento um exemplo de uma "mini-tabela" de roteamento:



Cada linha é uma entrada da tabela. Por exemplo, a linha a seguir é que define o Default Gateway da ser utilizado:

0.0.0.0 0.0.0.0 200.175.106.54 200.175.106.54 1

Uma entrada da tabela de roteamento possui os campos indicados no esquema a seguir e explicados logo em seguida:

Netwkor ID	Network Mask	Next Hop	Interface	Metric
0.0.0.0	0.0.0.0	200.175.106.54	200.175.106.54	1
10.100.100.0	255.255.255.0	10.200.200.4	10.200.200.4	1

- **Network ID:** Este é o endereço de destino. Pode ser o endereço de uma rede (por exemplo: 10.10.10.0), o endereço de um equipamento da rede, o endereço de uma sub-rede (veja detalhes sobre sub-redes na Parte 7) ou o endereço da rota padrão (0.0.0.0). A rota padrão significa: "a rota que será utilizada, caso não tenha sido encontrada uma rota específica para o destino". Por exemplo, se for definida que a rota padrão deve ser envida pela interface com IP 10.10.5.2 de um determinado roteador, sempre que chegar um pacote, para o qual não existe uma rota específica para o destino do pacote, este será enviado pela rota padrão, que no exemplo seria a interface 10.10.5.2. Falando de um jeito mais simples: Se não souber para onde mandar, manda para a rota padrão.
- **Network Mask**: A máscara de sub-rede utilizada para a rede de destino.

• **Next Hop:** Endereço IP da interface para a qual o pacote deve ser enviado. Considere o exemplo a seguir, como sendo uma entrada de um roteador, com uma interface de WAN configurada com o IP número 10.200.200.4:

Netwkor ID	Network Mask	Next Hop	Interface	Metric
10.100.100.0	255.255.255.0	10.200.200.1	10.200.200.120	1

Esta entrada indica que pacotes enviados para a rede definida pelos parâmetros 10.100.100.0/255.255.255.0, deve ser enviada para o gateway 10.200.200.1 e para chegar a este gateway, os pacotes de informação devem ser enviados pela interface 10.200.200.120. Neste exemplo, esta entrada está contida na tabela interna de roteamento de um computador com o Windows Server 2003, cujo número IP é 10.200.200.120 e o default gateway configurado é 10.200.200.1. Neste caso, quando este computador quiser se comunicar com um computador da rede 10.100.100.0, será usada a entrada de roteamento descrita neste item. Nesta entrada está especificado que pacotes para a rede 10.100.100.0, com máscara 255.255.255.0, devem ser enviados para o default gateway 10.200.200.1 e que este envio deve ser feito através da interface de rede 10.200.200.120, que no nosso exemplo é a placa de rede do computador. Uma vez que o pacote chegou no default gateway (na interface de LAN do roteador), o processo de roteamento, até a rede de destino (rede 10.100.100.0) é o processo descrito nas análises anteriores.

• **Interface:** É a interface através da qual o pacote deve ser enviado. Por exemplo, se você estiver analisando a tabela de roteamento interna, de um computador com o Windows Server 2003, o número IP do campo interface, será sempre o número IP da placa de rede, a não ser que você tenha mais de uma placa de rede instalada.

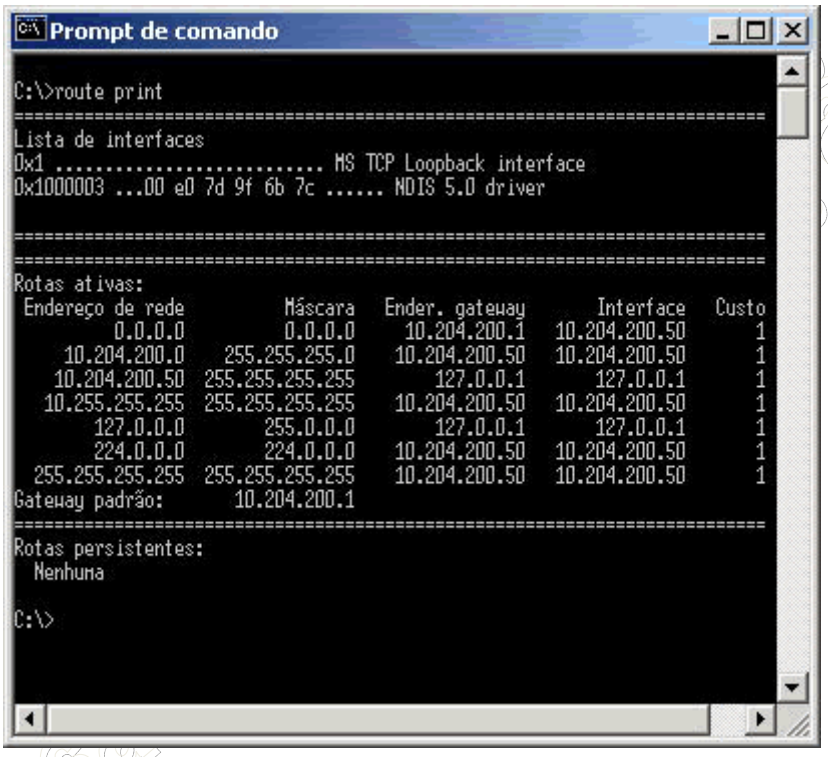
• **Metric:** A métrica é um indicativo da "distância" da rota, entre destino e origem, em termos de hopes. Conforme descrito anteriormente, pode haver mais de um roteador entre origem e destino. Também pode haver mais de um caminho entre origem e destino. Se for encontrada duas rotas para um mesmo destino, o roteamento será feito pela rota de menor valor no campo Metric. Um valor menor indica, normalmente, um número menor de hopes (roteadores) entre origem e destino.

Agora que você já conhece os conceitos de tabelas de roteamento e também conhece os campos que formam uma entrada em uma tabela de roteamento, é hora de analisar as entradas de uma tabela de roteamento em um computador com o Windows Server 2003 instalado. No Windows Server 2003, o protocolo TCP/IP é instalado automaticamente e não pode ser desinstalado (esta é uma das novidades do Windows Server 2003). Ao instalar e configurar o protocolo TCP/IP, o Windows Server 2003 cria, na memória do servidor, uma tabela de roteamento. Esta tabela é criada, dinamicamente, toda vez que o servidor é inicializado. Ao desligar o servidor o conteúdo desta tabela será descartado, para ser novamente recriado durante a próxima inicialização. A tabela de roteamento é criada com base nas configurações do protocolo TCP/IP. Existem também a possibilidade de adicionar entradas estáticas. Uma entrada estática fica gravada no HD do computador e será adicionada à tabela de roteamento durante a inicialização do sistema. Ou seja, além das entradas criadas automaticamente, com base nas configurações do TCP/IP, também podem ser acrescentadas rotas estáticas, criadas com o comando route, o qual descreverei mais adiante.

Para exibir a tabela de roteamento de um computador com o Windows Server 2003 (ou com o Windows 2000, ou Windows XP), abra um Prompt de comando (Iniciar -> Programas -> Acessórios -> Prompt de comando), digite o comando indicado a seguir e pressione Enter:

route print

Será exibida uma tabela de roteamento, semelhante a indicada na Figura a seguir, onde é exibida a tabela de roteamento para um servidor com o número IP: 10.204.200.50:



Vamos analisar cada uma destas entradas e explicar a função de cada entrada, para que você possa entender melhor os conceitos de roteamento.

Rota padrão

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
0.0.0.0	0.0.0.0	10.204.200.1	10.204.200.50	1

Esta rota é indicada por uma identificação de rede 0.0.0.0 com uma máscara de sub-rede 0.0.0.0. Quando o TCP/IP tenta encontrar uma rota para um determinado destino, ele percorre todas as entradas da tabela de roteamento em busca de uma rota específica para a rede de destino. Caso não seja encontrada uma rota para a rede de destino, será utilizada a rota padrão. Em outras palavras, se não houver uma rota específica, mande através da rota padrão. Observe que a rota padrão é justamente o default gateway da rede (10.204.200.1), ou seja, a interface de LAN do roteador da rede. O parâmetro Interface (10.204.200.50) é o número IP da placa de rede do próprio servidor. Em outras palavras: Se não houver uma rota específica manda para a rota padrão, onde o próximo hope da rede é o 10.204.200.1 e o envio para este hope é feito através da interface 10.204.200.50 (ou seja, a próprio placa de rede do servidor).

Endereço da rede local

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
10.204.200.0	255.255.255.0	10.204.200.50	10.204.200.50	1

Esta rota é conhecida como Rota da Rede Local. Ele basicamente diz o seguinte: "Quando o endereço IP de destino for um endereço da minha rede local, envie as informações através da minha placa

de rede (observe que tanto o parâmetro Gateway como o parâmetro Interface estão configurados com o número IP do próprio servidor). Ou seja, se for para uma das máquinas da minha rede local, manda através da placa de rede, não precisa enviar para o roteador.

Local host (endereço local)

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
10.204.200.50	255.255.255.255	127.0.0.1	127.0.0.1	1

Este endereço faz referência ao próprio computador. Observe que 10.204.200.50 é o número IP do servidor que está sendo analisado (no qual executei o comando route print). Esta rota diz que os programas do próprio computador, que enviarem pacotes para o destino 10.204.200.50 (ou seja, enviarem pacotes para si mesmo, como no exemplo de dois serviços trocando informações entre si), devem usar como Gateway o endereço de loopback 127.0.0.1, através da interface de loopback 127.0.0.1. Esta rota é utilizada para agilizar as comunicações que ocorrem entre os componentes do próprio Windows Server 2003, dentro do mesmo servidor. Ao usar a interface de loopback, toda a comunicação ocorre a nível de software, ou seja, não é necessário enviar o pacote através das diversas camadas do protocolo TCP/IP, até que o pacote chegue na camada de enlace (ou seja, a placa de rede), para depois voltar. Ao invés disso é utilizada a interface de loopback para direcionar os pacotes corretamente. Observe que esta entrada tem como máscara de sub-rede o número 255.255.255.255. Esta máscara indica que a entrada é uma rota para um endereço IP específico (no caso o próprio IP do servidor) e não uma rota para um endereço de rede.

Network broadcast (Broadcast de rede)

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
10.255.255.255	255.255.255.255	10.204.200.50	10.204.200.50	1

Esta rota define o endereço de broadcast da rede. Broadcast significa enviar para todos os computadores da rede. Quando é utilizado o endereço de broadcast, todos os computadores da rede recebem o pacote e processam o pacote. O broadcast é utilizado por uma série de serviços, como por exemplo o WINS, para fazer verificações periódicas de nomes, para enviar uma mensagem para todos os computadores da rede, para obter informações de todos os computadores e assim por diante. Observe que o gateway é o número IP da placa de rede do servidor e a Interface é este mesmo número, ou seja, para enviar um broadcast para a rede, envie através da placa de rede do servidor, não há necessidade de utilizar o roteador. Um detalhe interessante é que, por padrão, a maioria dos roteadores bloqueia o tráfego de broadcast, para evitar congestionamentos nos links de WAN.

Rede/endereço de loopback

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1

Comentei anteriormente que os endereços da rede 127.0.0.0 são endereços especiais, reservados para fazer referência a si mesmo. Ou seja, quando faço uma referência a 127.0.0.1 estou me referindo ao servidor no qual estou trabalhando. Esta roda indica, em palavras simples, que para se comunicar com a rede de loopback (127.0.0.0/255.0.0.0), utilize "eu mesmo" (127.0.0.1).

Multicast address (endereço de Multicast):

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
224.0.0.0	224.0.0.0	10.204.200.50	10.204.200.50	1

O tráfego IP, de uma maneira simples, pode ser de três tipos: Unicast é o tráfego direcionado para um número IP definido, ou seja, para um destinatário, definido por um número IP. Broadcast é o tráfego

dirigido para todos os computadores de uma ou mais redes. E tráfego Multicast é um tráfego direcionado para um grupo de computadores, os quais estão configurados e "inscritos" para receber o tráfego multicast. Um exemplo prático de utilização do multicast é para uma transmissão de vídeo através da rede. Vamos supor que de uma rede de 1000 computadores, apenas 30 devam receber um determinado arquivo de vídeo com um treinamento específico. Se for usado tráfego unicast, serão transmitidas 30 cópias do arquivo de vídeo (o qual já é um arquivo grande), uma cópia para cada destinatário. Com o uso do Multicast, uma única cópia é transmitida através do link de WAN e o tráfego multicast (com base no protocolo IGMP), entrega uma cópia do arquivo apenas para os 30 computadores devidamente configurados para receber o tráfego multicast. Esta rota define que o tráfego multicast deve ser enviado através da interface de rede, que é o número IP da placa de rede do servidor. Lembrando, quando falei sobre classes de endereços, a classe D é reservada para tráfego multicast, com IPs iniciando (o primeiro número) a partir de 224.

Limited Broadcast (Broadcast Limitado)

Endereço de rede	Máscara	Ender. gateway	Interface	Custo
255.255.255.255	255.255.255.255	10.204.200.50	10.204.200.50	1

Esta é a rota utilizada para o envio de broadcast limitado. O endereço de broadcast limitado é formato por todos os 32 bits do endereço IP sendo iguais a 1 (255.255.255.255). Este endereço é utilizado quando o computador tem que fazer o envio de um broadcast na rede local (envio do tipo um para todos na rede), porém o computador não conhece a número da rede local (network ID). Você pode perguntar: Mas em que situação o computador não conhecerá a identificação da rede local? Por exemplo, quando você inicializa um computador, configurado para obter as configurações do TCP/IP a partir de um servidor DHCP, a primeira coisa que este computador precisa fazer é localizar um servidor DHCP na rede e requisitar as configurações do TCP/IP. Ou seja, antes de receber as configurações do DHCP, o computador ainda não tem endereço IP e nem máscara de sub-rede, mas tem que se comunicar com um servidor DHCP. Esta comunicação é feita via broadcast limitado, onde o computador envia um pacote de formato específico (chamado de DHCP Discovery), para tentar descobrir um servidor DHCP na rede. Este pacote é enviado para todos os computadores. Aquele que for um servidor DHCP irá responder a requisição do cliente. Aí o processo de configuração do DHCP continua, até que o computador esteja com as configurações do TCP/IP definidas, configurações estas obtidas a partir do servidor DHCP.

9.5.4 Subnetting

Até agora, nas demais partes deste tutorial, sempre utilizei as máscaras de sub-rede padrão para cada classe de endereços, onde são utilizados oito, dezesseis ou vinte e quatro bits para a máscara de rede, conforme descrito a seguir:

Número de bits	Máscara de sub-rede
8	255.0.0.0
16	255.255.0.0
24	255.255.255.0

Por isso que existe uma outra notação, onde a máscara de sub-rede é indicada simplesmente pelo número de bits utilizados na máscara de sub-rede, conforme exemplos a seguir:

Definição da rede	Máscara de sub-rede
10.10.10.0/16	255.255.0.0
10.10.10.0/24	255.255.255.0
10.200.100.0/8	255.0.0.0

Porém com este esquema de endereçamento, baseado apenas nas máscaras de sub-rede padrão para cada classe (oito, dezesseis ou vinte e quatro bits), haveria um grande desperdício de números IP. Por exemplo, que empresa no mundo precisaria da faixa completa de uma rede classe A, na qual estão disponíveis mais de 16 milhões de endereços IP?

Vamos, agora, analisar o outro extremo desta questão. Imagine, por exemplo, uma empresa de porte médio, que tem a matriz em São Paulo e mais cinco filiais em outras cidades do Brasil. Agora imagine que em nenhuma das localidades, a rede tem mais do que 30 computadores. Se for usado as máscaras de sub-rede padrão, teria que ser definida uma rede Classe C (até 254 computadores), para cada localidade. Observe que estamos reservando 254 números IP para cada localidade (uma rede classe C com máscara 255.255.255.0), quando na verdade, no máximo, 30 números serão utilizados em cada localidade. Na prática, um belo desperdício de endereços IP, mesmo em uma empresa de porte médio ou pequeno.

Observe que neste exemplo, uma única rede Classe C seria suficiente. Já que são seis localidades (a matriz mais seis filiais), com um máximo de 30 endereços por localidade, um total de 254 endereços de uma rede Classe C seria mais do que suficiente. Ainda haveria desperdício, mas agora bem menor.

A boa notícia é que é possível “dividir” uma rede (qualquer rede) em sub-redes, onde cada sub-rede fica apenas com uma faixa de números IP de toda a faixa original. Por exemplo, a rede Classe C 200.100.100.0/255.255.255.0, com 256 números IPs disponíveis (na prática são 254 números que podem ser utilizados, descontando o primeiro que é o número da própria rede e o último que o endereço de broadcast), poderia ser dividida em 8 sub-redes, com 32 números IP em cada sub-rede. O esquema a seguir ilustra este conceito:

Rede original: 256 endereços IP disponíveis: 200.100.100.0 -> 200.100.100.255

Divisão da rede em 8 sub-redes, onde cada sub-rede fica com 32 endereços IP:

Sub-rede 01: 200.100.100.0 -> 200.100.100.31
Sub-rede 02: 200.100.100.32 -> 200.100.100.63
Sub-rede 03: 200.100.100.64 -> 200.100.100.95
Sub-rede 04: 200.100.100.96 -> 200.100.100.127
Sub-rede 05: 200.100.100.128 -> 200.100.100.159
Sub-rede 06: 200.100.100.160 -> 200.100.100.191
Sub-rede 07: 200.100.100.192 -> 200.100.100.223
Sub-rede 08: 200.100.100.224 -> 200.100.100.255

Para o exemplo da empresa com seis localidades (matriz mais cinco filiais), onde, no máximo, são necessários trinta endereços IP por localidade, a utilização de uma única rede classe C, dividida em 8 sub-redes seria a solução ideal. Na prática a primeira e a última sub-rede são descartadas, pois o primeiro IP da primeira sub-rede representa o endereço de rede e o último IP da última sub-rede representa o endereço de broadcast. Com isso restariam, ainda, seis sub-redes. Exatamente a quantia necessária para o exemplo proposto. Observe que ao invés de seis redes classe C, bastou uma única rede Classe C, subdividida em seis sub-redes. Uma bela economia de endereços. Claro que se um dos escritórios, ou a matriz, precisasse de mais de 32 endereços IP, um esquema diferente de divisão teria que ser criado.

Entendido o conceito teórico de divisão em sub-redes, resta o trabalho prático, ou seja:

- O que tem que ser alterado para fazer a divisão em sub-redes (sub netting)?
- Como calcular o número de sub-redes e o número de números IP dentro de cada sub-rede?
- Como listar as faixas de endereços dentro de cada sub-rede?
- Exemplos práticos

Você aprenderá estas etapas através de exemplos práticos. Vou inicialmente mostrar o que tem que ser alterado para fazer a divisão de uma rede padrão (com máscara de 8, 16 ou 24 bits) em uma ou mais sub-redes. Em seguida, apresento alguns exemplos de divisão de uma rede em sub-redes. Mãos a obra.

Por padrão são utilizadas máscaras de sub-rede de 8, 16 ou 24 bits, conforme indicado no esquema a seguir:

Número de bits	Máscara de sub-rede
08	255.0.0.0
16	255.255.0.0
24	255.255.255.0

Uma máscara de 8 bits significa que todos os bits do primeiro octeto são iguais a 1; uma máscara de 16 bits significa que todos os bits do primeiro e do segundo octeto são iguais a 1 e uma máscara de 24 bits significa que todos os bits dos três primeiros octetos são iguais a 1. Este conceito está ilustrado na tabela a seguir:

Núm. bits	Octeto 01	Octeto 02	Octeto 03	Octeto 04	Máscara
8	11111111	00000000	00000000	00000000	255.0.0.0
16	11111111	11111111	00000000	00000000	255.255.0.0
24	11111111	11111111	11111111	00000000	255.255.255.0

No exemplo da rede com matriz em São Paulo e mais cinco escritórios, vamos utilizar uma rede classe C, que será subdividida em seis sub-redes (na prática 8, mas a primeira e a última não são utilizadas). Para fazer esta subdivisão, você deve alterar o número de bits iguais a 1 na máscara de sub-rede. Por exemplo, ao invés de 24 bits, você terá que utilizar 25, 26, 27 ou um número a ser definido. Bem, já avançamos mais um pouco:

“Para fazer a divisão de uma rede em sub-redes, é preciso aumentar o número de bits iguais a 1, alterando com isso a máscara de sub-rede.”

Quantos bits devem ser utilizados para a máscara de sub-rede?

Agora, naturalmente, surge uma nova questão: "Quantos bits?". Ou de uma outra maneira (já procurando induzir o seu raciocínio): "O que define o número de bits a ser utilizados a mais?"

Bem, esta é uma questão bem mais simples do que pode parecer. Vamos a ela. No exemplo proposto, precisamos dividir a rede em seis sub-redes. Ou seja, o número de sub-redes deve ser, pelo menos, seis. Sempre lembrando que a primeira e a última sub-rede não são utilizadas. O número de sub-redes é proporcional ao número de bits que vamos adicionar à máscara de sub-rede já existente. O número de rede é dado pela fórmula a seguir, onde 'n' é o número de bits a mais a serem utilizados para a máscara de sub-rede:

$$\text{Núm. de sub-redes} = 2^n-2$$

No nosso exemplo estão disponíveis até 8 bits do último octeto para serem também utilizados na máscara de sub-rede. Claro que na prática não podemos usar os 8 bits, senão ficaríamos com o endereço de broadcast: 255.255.255.255, como máscara de sub-rede. Além disso, quanto mais bits eu pegar para a máscara de sub-rede, menos sobrarão para os números IP da rede. Por exemplo, se eu adicionar mais um bit a máscara já existente, ficarei com 25 bits para a máscara e 7 para números IP, se eu adicionar mais dois bits à máscara original de 24 bits, ficarei com 26 bits para a máscara e somente 6 para números IP e assim por diante. O número de bits que restam para os números IP, definem quantos números IP podem haver em cada sub-rede. A fórmula para determinar o número de endereços IP dentro de cada sub-rede, é indicado a seguir, onde 'n' é o núméo de bits destinados a parte de host do endereço (32 – bits usados para a máscara):

$$\text{Núm. de end. IP dentro de cada sub-rede} = 2^n-2$$

Na tabela a seguir, apresento cálculos para a divisão de sub-redes que será feita no nosso exemplo. Observe que quanto mais bits eu adiciono à máscara de sub-rede, mais sub-redes é possível obter, porém com um menor número de máquinas em cada sub-rede. Lembrando que o nosso exemplo estamos subdividindo uma rede classe C - **200.100.100.0/255.255.255.0**, ou seja, uma rede com 24 bits para a máscara de sub-rede original.

Número de bits a mais a serem utilizados	Número de sub-redes	Número de hosts em cada sub-rede
0	máscara original. rede classe C sem divisão	254
1	0	126
2	2	62
3	6	30
4	14	14
5	30	6
6	62	2
7	126	0
8	endereço de broadcast	-

Número de redes e número de hosts em cada rede

Claro que algumas situações não se aplicam na prática. Por exemplo, usando apenas um bit a mais para a máscara de sub-rede, isto é, 25 bits ao invés de 24. Neste caso teremos 0 sub-redes disponíveis. Pois com 1 bit é possível criar apenas duas sub-redes, como a primeira e a última são descartadas, conforme descrito anteriormente, na prática as duas sub-redes geradas não poderão ser utilizadas. A mesma situação ocorre com o uso de 7 bits a mais para a máscara de sub-rede, ou seja, 31 ao invés de 24. Nesta situação sobra apenas um bit para os endereços IP. Com 1 bit posso ter apenas dois endereços IP, descontanto o primeiro e o último que não são utilizados, não sobra nenhum endereço IP. As situações intermediárias é que são mais realistas. No nosso exemplo, precisamos dividir a rede Classe C - 200.100.100.0/255.255.255.0, em seis sub-redes. De acordo com a tabela da Figura anterior, precisamos utilizar 3 bits a mais para obter as seis sub-redes desejadas.

Observe que utilizando três bits a mais, ao invés de 24 bits (máscara original), vamos utilizar 27 bits para a máscara de sub-rede. Com isso sobra cinco bits para os números IPs dentro de cada sub-rede, o que dá um total de 30 números IP por sub-rede. Exatamente o que precisamos.

A próxima questão que pode surgir é como é que fica a máscara de sub-rede, agora que ao invés de 24 bits, estou utilizando 27 bits, conforme ilustrado na tabela a seguir:

Núm. bits	Octeto 01	Octeto 02	Octeto 03	Octeto 04
27	11111111	11111111	11111111	11100000

Para determinar a nova máscara temos que revisar o valor de cada bit. Da esquerda para a direita, cada bit representa o seguinte valor, respectivamente:

128 64 32 16 8 4 2 1

Como os três primeiros bits do último octeto foram também utilizados para a máscara, estes três bits soman para o valor do último octeto. No nosso exemplo, o último octeto da máscara terá o seguinte valor: 128+64+32 = 224. Com isso a nova máscara de sub-rede, máscara esta que será utilizada pelas seis sub-redes, é a seguinte: 255.255.255.224. Observe que ao adicionarmos bits à máscara de sub-rede, fazemos isso a partir do bit de maior valor, ou seja, o bit mais da esquerda, com o valor de 128, depois usamos o próximo bit com valor 64 e assim por diante. Na tabela a seguir, apresento a ilustração de como fica a nova máscara de sub-rede:

Núm. bits	Octeto 01	Octeto 02	Octeto 03	Octeto 04	Nova Máscara
27	11111111	11111111	11111111	11100000	255.255.255.224

Com o uso de três bits adicionais para a máscara de rede, teremos seis sub-redes disponíveis (uma para cada escritório) com um número máximo de 30 números IP por sub-rede. Exatamente o que precisamos para o exemplo proposto. A idéia básica de subnetting é bastante simples. Utiliza-se bits adicionais para a máscara de sub-rede. Com isso tenho uma divisão da rede original (classe A, classe B ou classe C) em várias sub-redes, sendo que o número de endereços IP em cada sub-rede é reduzido (por termos utilizados bits adicionais para a máscara de sub-rede, bits estes que originalmente eram destinados aos endereços IP). Esta divisão pode ser feita em redes de qualquer uma das classes padrão A, B ou C. Por exemplo, por padrão, na Classe A são utilizados 8 bits para a máscara de sub-rede e 24 bits para hosts. Você pode utilizar, por exemplo, 12 bits para a máscara de sub-rede, restando com isso 20 bits para endereços de host.

Na tabela a seguir, apresento os cálculos para o número de sub-redes e o número de hosts dentro de cada sub-rede, apenas para os casos que podem ser utilizados na prática, ou seja, duas ou mais sub-redes e dois ou mais endereços válidos em cada sub-rede, quando for feita a sub-divisão de uma rede Classe C, com máscara original igual a 255.255.255.0..

Número de bits a mais a serem utilizados	Número de sub-redes	Número de hosts em cada sub-rede
2	2	62
3	6	30
4	14	14
5	30	6
6	62	2

Lembrando que a fórmula para calcular o número de sub-redes é:

Núm. de sub-redes = 2^n-2

onde n é o número de bits a mais utilizados para a máscara de sub-rede

E a fórmula para calcular o número de endereços IP dentro de cada sub-rede é:

2^n-2

onde n é o número de bits restantes, isto é, não utilizados pela máscara de sub-rede.

Até aqui trabalhei com um exemplo de uma rede Classe C, que está sendo subdividida em várias sub-redes. Porém é também possível subdividir redes Classe A e redes Classe B. Lembrando que redes classe A utilizam, por padrão, apenas 8 bits para o endereço de rede, já redes classe B, utilizam, por padrão, 16 bits. Na tabela a seguir, apresento um resumo do número de bits utilizados para a máscara de sub-rede, por padrão, nas classes A, B e C:

Classe	Bits	Octeto 01	Octeto 02	Octeto 03	Octeto 04	Máscara padrão
A	8	11111111	00000000	00000000	00000000	255.0.0.0
B	16	11111111	11111111	00000000	00000000	255.255.0.0
C	24	11111111	11111111	11111111	00000000	255.255.255.0

Para subdividir uma rede classe A em sub-redes, basta usar bits adicionais para a máscara de sub-rede. Por padrão são utilizados 8 bits. Se você utilizar 10, 12 ou mais bits, estará criando sub-redes. O mesmo raciocínio é válido para as redes classe B, as quais utilizam, por padrão, 16 bits para a máscara de sub-rede. Se você utilizar 18, 20 ou mais bits para a máscara de sub-rede, estará subdividindo a rede classe B em várias sub-redes.

As fórmulas para cálculo do número de sub-redes e do número de hosts em cada sub-rede são as mesmas apresentadas anteriormente, independentemente da classe da rede que está sendo dividida em sub-redes.

A seguir apresento uma tabela com o número de sub-redes e o número de hosts em cada sub-rede, dependendo do número de bits adicionais (além do padrão definido para a classe) utilizados para a máscara de sub-rede, para a divisão de uma rede Classe B:

Divisão de uma rede classe B em sub-redes			
Número de bits	Sub-redes	Hosts	Nova máscara de sub-rede
2	2	16382	255.255.192.0
3	6	8190	255.255.224.0
4	14	4094	255.255.240.0
5	30	2046	255.255.248.0
6	62	1022	255.255.252.0
7	126	510	255.255.254.0
8	254	254	255.255.255.0
9	510	126	255.255.255.128
10	1022	62	255.255.255.192
11	2046	30	255.255.255.224
12	4094	14	255.255.255.240
13	8190	6	255.255.255.248

Observe como o entendimento dos cálculos binários realizados pelo TCP/IP facilita o entendimento de vários assuntos relacionados ao TCP/IP, inclusive o conceito de subnetting (Veja Parte 2 para detalhes sobre Cálculos Binários). Por padrão a classe B utiliza 16 bits para a máscara de sub-rede, ou seja, uma máscara padrão: 255.255..0.0. Agora se utilizarmos oito bits adicionais (todo o terceiro octeto) para a máscara, teremos todos os bits do terceiro octeto como sendo iguais a 1, com isso a máscara passa a ser: 255.255.255.0. Este resultado está coerente com a tabela da Figura 16.11. Agora vamos avançar um pouco mais. Ao invés de 8 bits adicionais, vamos utilizar 9. Ou seja, todo o terceiro octeto (8 bits) mais o primeiro bit do quarto octeto. O primeiro bit, o bit bem à esquerda é o bit de valor mais alto, ou seja, o que vale 128. Ao usar este bit também para a máscara de sub-rede, obtemos a seguinte máscara: 255.255.255.128. Também fecha com a tabela anterior. Com isso você pode concluir que o entendimento da aritemética e da representação binária, facilita muito o estudo do protocolo TCP/IP e de assuntos relacionados, tais como subnetting e roteamento.

A seguir apresento uma tabela com o número de sub-redes e o número de hosts em cada sub-rede, dependendo do número de bits adicionais (além do padrão definido para a classe) utilizados para a máscara de sub-rede, para a divisão de uma rede Classe A:

Divisão de uma rede classe C em sub-redes			
Número de bits	Sub-redes	Hosts	Nova máscara de sub-rede
2	2	4194302	255.192.0.0
3	6	2097150	255.224.0.0
4	14	1048574	255.240.0.0
5	30	524286	255.248.0.0
6	62	262142	255.252.0.0
7	126	131070	255.254.0.0
8	254	65534	255.255.0.0
9	510	32766	255.255.128.0
10	1022	16382	255.255.192.0
11	2046	8190	255.255.224.0
12	4094	4094	255.255.240.0
13	8190	2046	255.255.248.0
14	16382	1022	255.255.252.0
15	32766	510	255.255.254.0
16	65534	254	255.255.255.0
17	131070	126	255.255.255.128
18	262142	62	255.255.255.192
19	524286	30	255.255.255.224
20	1048574	14	255.255.255.240
21	2097150	6	255.255.255.248
22	4194302	2	255.255.255.252

Um fato importante, que eu gostaria de destacar novamente é que todas as sub-redes (resultantes da divisão de uma rede), utilizam o mesmo número para a máscara de sub-rede. Por exemplo, na quarta linha da tabela indicada na Figura 16.12, estou utilizando 5 bits adicionais para a máscara de sub-rede, o que resulta em 30 sub-redes diferentes, porém todas utilizando como máscara de sub-rede o seguinte número: 255.248.0.0.

Muito bem, entendido o conceito de divisão em sub-redes e de determinação do número de sub-redes, do número de hosts em cada sub-rede e de como é formada a nova máscara de sub-rede, a próxima questão que pode surgir é a seguinte:

Como listar as faixas de endereços para cada sub-rede? Este é exatamente o assunto que vem a seguir. Vamos entender esta questão através de exemplos práticos.

Exemplo 01: Dividir a seguinte rede classe C: 229.45.32.0/255.255.255.0. São necessárias, pelo menos, 10 sub-redes. Determinar o seguinte:

a) Quantos bits serão necessários para fazer a divisão e obter pelo menos 10 sub-redes?

b) Quantos números IP (hosts) estarão disponíveis em cada sub-rede?

c) Qual a nova máscara de sub-rede?

d) Listar a faixa de endereços de cada sub-rede.

Vamos ao trabalho. Para responder a questão da letra a, você deve lembrar da fórmula:

$$\textbf{Núm. de sub-redes} = 2^n - 2$$

Você pode ir substituindo n por valores sucessivos, até atingir ou superar o valor de 10. Por exemplo, para n=2, a fórmula resulta em 2, para n=3, a fórmula resulta em 6, para n=4 a fórmula resulta em 14. Bem, está respondida a questão da letra a, temos que utilizar quatro bits do quarto octeto para fazer parte da máscara de sub-rede.

a) Quantos bits serão necessários para fazer a divisão e obter pelo menos 10 sub-redes?

R: 4 bits.

Como utilizei quatro bits do último octeto (além dos 24 bits dos três primeiros octetos, os quais já faziam parte da máscara original), sobraram apenas 4 bits para os endereços IP, ou seja, para os endereços de hosts em cada sub-rede. Tenho que lembrar da seguinte fórmula:

$$\textbf{Núm. de end. IP dentro de cada sub-rede} = 2^n - 2$$

substituindo n por 4, vou obter um valor de 14. Com isso já estou em condições de responder a alternativa b.

b) Quantos números IP (hosts) estarão disponíveis em cada sub-rede?

R: 14.

Como utilizei quatro bits do quarto octeto para fazer a divisão em sub-redes, os quatro primeiros bits foram definidos iguais a 1. Basta somar os respectivos valores, ou seja: 128+64+32+16 = 240. Ou seja, com os quatro primeiros bits do quarto octeto sendo iguais a 1, o valor do quarto octeto passa para 240, com isso já temos condições de responder a alternativa c.

c) Qual a nova máscara de sub-rede?

R: 255.255.255.240

É importante lembrar, mais uma vez, que esta será a máscara de sub-rede utilizada por todas as 14 sub-redes.

d) Listar a faixa de endereços de cada sub-rede.

Esta é a novidade deste item. Como saber de que número até que número vai cada endereço IP. Esta também é fácil, embora seja novidade. Observe o último bit definido para a máscara. No nosso exemplo é o quarto bit do quarto octeto. Qual o valor decimal do quarto bit? 16 (o primeiro é 128, o segundo 64, o terceiro 32 e assim por diante, conforme explicado na Parte 2). O valor do último bit é um indicativo das faixas de variação para este exemplo. Ou seja, na prática temos 16 hosts em cada sub-rede, embora o primeiro e o último não devam ser utilizados, pois o primeiro é o endereço da própria sub-rede e o último é o endereço de broadcast da sub-rede. Por isso que ficam 14 hosts por sub-rede, devido ao '-2' na fórmula, o '-2' significa: - o primeiro – o último. Ao listar as faixas, consideramos os 16 hosts, apenas é importante salienar que o primeiro e o último não são utilizados. Com isso a primeira sub-rede vai do host 0 até o 15, a segunda sub-rede do 16 até o 31, a terceira do 32 até o 47 e assim por diante, conforme indicado no esquema a seguir:

Divisão da rede em 14 sub-redes, onde cada sub-rede fica com 16 endereços IP, sendo que a primeira e a última sub-rede não são utilizadas e o primeiro e o último número IP, dentro de cada sub-rede, também não são utilizados:

```
Sub-rede 01 229.45.32.0      ->   229.45.32.15
Sub-rede 02  229.45.32.16    ->   229.45.32.31
Sub-rede 03  229.45.32.32    ->   229.45.32.47
Sub-rede 04  229.45.32.48    ->   229.45.32.63
Sub-rede 05  229.45.32.64    ->   229.45.32.79
Sub-rede 06  229.45.32.80    ->   229.45.32.95
Sub-rede 07  229.45.32.96    ->   229.45.32.111
Sub-rede 08 229.45.32.112    ->   229.45.32.127
Sub-rede 09 229.45.32.128    ->   229.45.32.143
Sub-rede 10  229.45.32.144   ->   229.45.32.159
Sub-rede 11  229.45.32.160   ->   229.45.32.175
Sub-rede 12  229.45.32.176   ->   229.45.32.191
Sub-rede 13  229.45.32.192   ->   229.45.32.207
Sub-rede 14  229.45.32.208   ->   229.45.32.223
Sub-rede 15  229.45.32.224   ->   229.45.32.239
Sub-rede 16  229.45.32.240   ->   229.45.32.255
```

Vamos a mais um exemplo prático, agora usando uma rede classe B, que tem inicialmente, uma máscara de sub-rede: 255.255.0.0

Exemplo 02: Dividir a seguinte rede classe B: 150.100.0.0/255.255.0.0. São necessárias, pelo menos, 20 sub-redes. Determinar o seguinte:

a) Quantos bits serão necessários para fazer a divisão e obter pelo menos 10 sub-redes?

b) Quantos números IP (hosts) estarão disponíveis em cada sub-rede?

c) Qual a nova máscara de sub-rede?

d) Listar a faixa de endereços de cada sub-rede.

Vamos ao trabalho. Para responder a questão da letra a, você deve lembrar da fórmula:

$$\textbf{Núm. de sub-redes} = 2^n - 2$$

Você pode ir substituindo n por valores sucessivos, até atingir ou superar o valor de 10. Por exemplo, para n=2, a fórmula resulta em 2, para n=3, a fórmula resulta em 6, para n=4 a fórmula resulta em 14 e para n=5 a fórmula resulta em 30. Bem, está respondida a questão da letra a, temos que utilizar cinco bits do terceiro octeto para fazer parte da máscara de sub-rede. Pois se utilizarmos apenas 4 bits, obteremos

somente 14 sub-redes e usando mais de 5 bits, obteremos um número de sub-redes bem maior do que o necessário.

a) Quantos bits serão necessários para fazer a divisão e obter pelo menos 20 sub-redes?

R: 5 bits.

Como utilizei cinco bits do terceiro octeto (além dos 16 bits dos dois primeiros octetos, os quais já faziam parte da máscara original)., sobraram apenas 11 bits (os três restantes do terceiro octeto mais os 8 bits do quarto octeto) para os endereços IP, ou seja, para os endereços de hosts em cada sub-rede. Tenho que lembrar da seguinte fórmula:

Núm. de endereços IP dentro de cada sub-rede = 2^n-2

substituindo n por 11 (número de bits que restarama para a parte de host), vou obter um valor de 2046, já descontando o primeiro e o último número, os quais não podem ser utilizados, conforme já descrito anteriormente. Com isso já estou em condições de responder a alternativa b.

b) Quantos números IP (hosts) estarão disponíveis em cada sub-rede?

R: 2046.

Como utilizei cinco bits do terceiro octeto para fazer a divisão em sub-redes, os cinco primeiros bits foram definidos iguais a 1. Basta somar os respectivos valores, ou seja: 128+64+32+16+8 = 248. Ou seja, com os quatro primeiros bits do quarto octeto sendo iguais a 1, o valor do quarto octeto passa para 248, com isso já temos condições de responder a alternativa c.

c) Qual a nova máscara de sub-rede?

R: 255.255.248.0

É importante lembrar, mais uma vez, que esta será a máscara de sub-rede utilizada por todas as 30 sub-redes.

d) Listar a faixa de endereços de cada sub-rede.

Como saber de que número até que número vai cada endereço IP. Esta também é fácil e o raciocínio é o mesmo utilizado para o exemplo anterior, onde foi feita uma divisão de uma rede classe C. Observe o último bit definido para a máscara. No nosso exemplo é o quinto bit do terceiro octeto. Qual o valor decimal do quinto bit (de qualque octeto)? 8 (o primeiro é 128, o segundo 64, o terceiro 32, o quarto é 16 e o quinto é 8). O valor do último bit é um indicativo das faixas de variação para este exemplo. Ou seja, na prática temos 2048 hosts em cada sub-rede, embora o primeiro e o último não devam ser utilizados, pois o primeiro é o endereço da própria sub-rede e o último é o endereço de broadcast da sub-rede. Por isso que ficam 2046 hosts por sub-rede, devido ao '-2' na fórmula, o '-2' significa: - o primeiro – o último. Ao listar as faixas, consideramos o valor do último bit da máscara. No nosso exemplo é o 8. A primeira faixa vai do zero até um número anterior ao valor do último bit, no caso do 0 ao 7. A seguir indico a faixa de endereços da primeira sub-rede (sub-rede que não será utilizada na prática, pois descarta-se a primeira e a última):

Sub-rede 01 150.100.0.1 -> 150.100.7.254

Com isso todo endereço IP que tiver o terceiro número na faixa entre 0 e 7, será um número IP da primeira sub-rede, conforme os exemplos a seguir: 150.100.0.25, 150.100.3.20, 150.100.5.0, 150.100.6.244

Importante: Observe que os valores de 0 a 7 são definidos no terceiro octeto, que é onde estamos utilizando cinco bits a mais para fazer a divisão em sub-redes.

Qual seria a faixa de endereços IP da próxima sub-rede. Aqui vale o mesmo raciocínio. O último bit da máscara equivale ao valor 8. Esta é a variação da terceira parte do número IP, que é onde esta sendo

feita a divisão em sub-redes. Então, se a primeira foi de 0 até 7, a segunda sub-rede terá valores de 8 a 15 no terceiro octeto, a terceira sub-rede terá valores de 16 a 23 e assim por diante.

Divisão da rede em 32 sub-redes, onde cada sub-rede fica com 2048 endereços IP, sendo que a primeira e a última sub-rede não são utilizadas e o primeiro e o último número IP, dentro de cada sub-rede, também não são utilizados:

Sub-rede	Primeiro IP	Último IP	End. de broadcast	Número
150.100.0.0	150.100.0.1	150.100.7.254	150.100.7.255	01
150.100.8.0	150.100.8.1	150.100.15.254	150.100.15.255	02
150.100.16.0	150.100.16.1	150.100.23.254	150.100.23.255	03
150.100.24.0	150.100.24.1	150.100.31.254	150.100.31.255	04
150.100.32.0	150.100.32.1	150.100.39.254	150.100.39.255	05
150.100.40.0	150.100.40.1	150.100.47.254	150.100.47.255	06
150.100.48.0	150.100.48.1	150.100.55.254	150.100.55.255	07
150.100.56.0	150.100.56.1	150.100.63.254	150.100.63.255	08
150.100.64.0	150.100.64.1	150.100.71.254	150.100.71.255	09
150.100.72.0	150.100.72.1	150.100.79.254	150.100.79.255	10
150.100.80.0	150.100.80.1	150.100.87.254	150.100.87.255	11
150.100.88.0	150.100.88.1	150.100.95.254	150.100.95.255	12
150.100.96.0	150.100.96.1	150.100.103.254	150.100.103.255	13
150.100.104.0	150.100.104.1	150.100.111.254	150.100.111.255	14
150.100.112.0	150.100.112.1	150.100.119.254	150.100.119.255	15
150.100.120.0	150.100.120.1	150.100.127.254	150.100.127.255	16
150.100.128.0	150.100.128.1	150.100.135.254	150.100.135.255	17
150.100.136.0	150.100.136.1	150.100.143.254	150.100.143.255	18
150.100.144.0	150.100.144.1	150.100.151.254	150.100.151.255	19
150.100.152.0	150.100.152.1	150.100.159.254	150.100.159.255	20
150.100.160.0	150.100.160.1	150.100.167.254	150.100.167.255	21
150.100.168.0	150.100.168.1	150.100.175.254	150.100.175.255	22
150.100.176.0	150.100.176.1	150.100.183.254	150.100.183.255	23
150.100.184.0	150.100.184.1	150.100.191.254	150.100.191.255	24
150.100.192.0	150.100.192.1	150.100.199.254	150.100.199.255	25
150.100.200.0	150.100.200.1	150.100.207.254	150.100.207.255	26
150.100.208.0	150.100.208.1	150.100.215.254	150.100.215.255	27
150.100.216.0	150.100.216.1	150.100.223.254	150.100.223.255	28
150.100.224.0	150.100.224.1	150.100.231.254	150.100.231.255	29
150.100.232.0	150.100.232.1	150.100.239.254	150.100.239.255	30
150.100.240.0	150.100.240.1	150.100.247.254	150.100.247.255	31
150.100.248.0	150.100.248.1	150.100.255.254	150.100.255.255	32

Com base na tabela apresentada, fica fácil responder em que sub-rede está contido um determinado número IP. Por exemplo, considere o número IP 1500.100.130.222. Primeiro você observa o terceiro octeto do número IP (o terceiro, porque é neste octeto que estão os últimos bits que foram utilizados para a máscara de sub-rede). Consultando a tabela anterior, você observa o valor de 130 para o terceiro octeto corresponde a sub-rede 17, na qual o terceiro octeto varia entre 128 e 135, conforme indicado a seguir:

150.100.128.0 150.100.128.1 150.100.135.254 150.100.135.255 17

9.6 NÍVEL DE TRANSPORTE

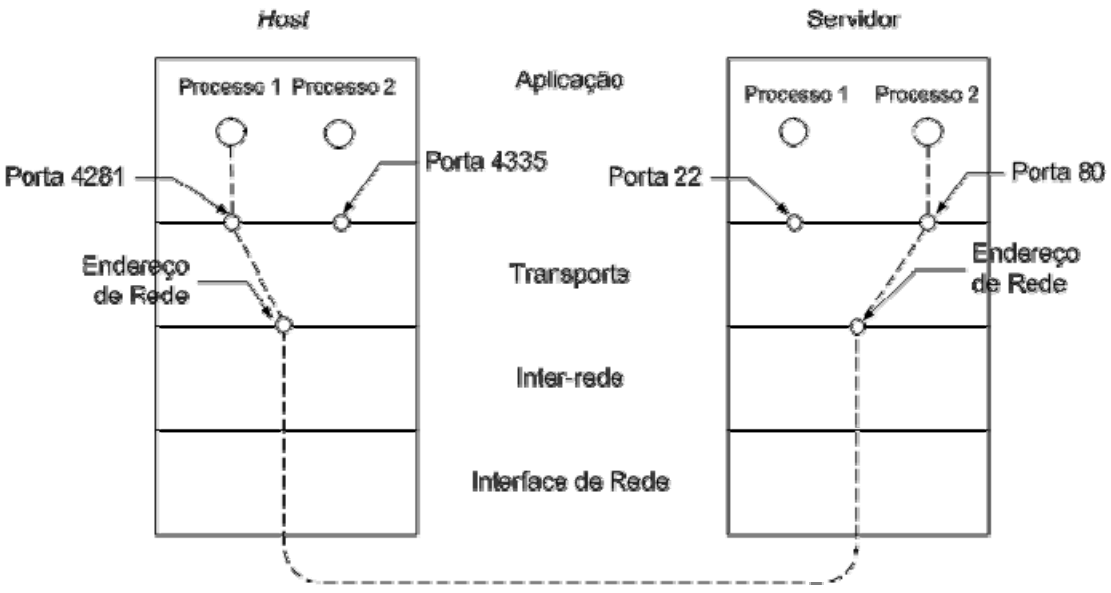
A função básica da camada de transporte é aceitar dados da camada de aplicação, dividi-los em unidades menores em caso de necessidade, passá-los para a camada de rede e garantir que todas essas unidades cheguem corretamente à outra extremidade. Além disso, tudo tem de ser feito com eficiência de forma que as camadas superiores fiquem isoladas das inevitáveis mudanças na tecnologia de hardware.

A camada de transporte é uma camada fim a fim, que liga a origem ao destino. Em outras palavras, um programa da máquina de origem mantém uma conversa com um programa semelhante instalado na máquina de destino, utilizando cabeçalhos de mensagem e mensagens de controle. Entre as camadas de transporte de diferentes hosts são trocadas TPDUs (Transport Protocol Data Units) chamados de segmentos. Um segmento é composto pelo cabeçalho da camada de transporte e os dados da camada de aplicação.

Muitos hosts são multiprogramados; isso significa que muitas conexões estarão entrando e saindo de cada host. Desta forma é responsabilidade da camada de transporte multiplexar todas as comunicações em um único canal determinando a qual conexão uma mensagem pertence. Além da multiplexação é responsabilidade desta camada estabelecer conexões, encerrá-las e controlá-las de forma que um host muito rápido não possa sobrecarregar um host muito lento (controle de fluxo). Em redes IP são utilizados dois protocolos para a implementação destas funções: o TCP e o UDP.

Da mesma forma que em outras camadas, a camada de transporte também possui um endereçamento. Quando um processo de aplicação deseja estabelecer uma conexão com um processo de aplicação remoto, é necessário especificar a aplicação com a qual ele irá se conectar. O método utilizado é definir os endereços de transporte que os processos podem ouvir para receber solicitações de conexão.

Os processos utilizam os TSAP (Transport Service Acess Point – Ponto de Acesso de Serviços de Transporte) para se intercomunicarem. Em redes IP, o TSAP é um número de 16 bits chamado de porta. O endereço da camada de transporte é um número de 48 bits, composto pela agregação do endereço IP do host e o número da porta. Os serviços da camada de transporte são obtidos através da comunicação entre os sockets do transmissor e do receptor.



Se calcularmos (2^16) veremos que existem ao todo 65536 portas a serem utilizadas. Para uma melhor organização de serviços, algumas portas foram definidas pela IANA (**I**nternet **A**ssigned **N**umbers **A**uthority) como “portas bem conhecidas” (well-known ports). Estas são as portas abaixo de 1024, para aplicações não padronizadas são utilizadas portas acima deste valor.

Dessa forma, para cada protocolo da camadas de transporte, temos 65536 portas para o TCP, 65536 portas para o UDP, 65536 portas... Conhecer essas portas é fundamental para operar um Firewall de forma satisfatória.

"Ok, mas com tanta porta como vou poder saber todas elas?". Bem, você não precisa conhecer todas. Até mesmo porque a maior parte delas não são especificadas. Para a nossa alegria, apenas as primeiras 1024 são especificadas. Acho que não ajudou muito, ne?! Ok, vamos melhorar. Para um administrador de rede é imprecindível saber pelomenos as portas dos serviços básicos de Rede: telnet, SSh, FTP, SMTP, POP, HTTP, HTTPS... Não são muitas, mas antes de ver isso, vamos entender que controla essas portas.

Abaixo a lista de algumas das principais portas TCP, e no site: http://www.iana.org/assignments/port-numbers, a listagem completa:

serviço	porta	protocolo
daytime	13	tcp e udp
ftp-data	20	tcp
ftp	21	tcp
ssh	22	tcp
telnet	23	tcp
smtp	25	tcp e udp
name	42	tcp e udp
nameserver	42	tcp e udp
tftp	69	tcp e udp
www	80	tcp
pop3	110	tcp e udp
netbios-ns	137	tcp e udp
netbios-dgm	138	tcp e udp
netbios-ssn	139	tcp e udp

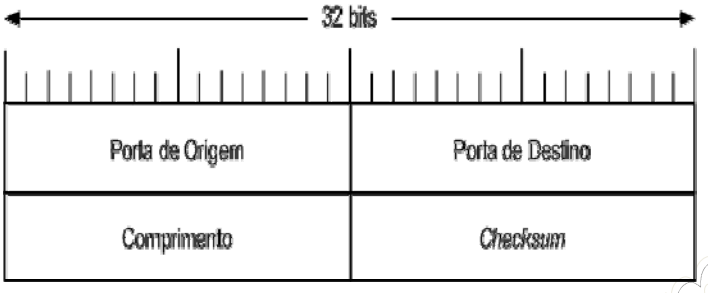
Os sockets são diferentes para cada protocolo de transporte, desta forma mesmo que um socket TCP possua o mesmo número que um socket UDP, ambos são responsáveis por aplicações diferentes. Os sockets de origem e destino são responsáveis pela identificação única da comunicação. Desta forma é possível a implementação da função conhecida como multiplexação. A multiplexação possibilita que haja várias conexões partindo de um único host ou terminando em um mesmo servidor.

A formação do socket de origem e destino se dá da seguinte forma:

1. Ao iniciar uma comunicação é especificado para a aplicação o endereço IP de destino e a porta de destino;
2. A porta de origem é atribuída dinamicamente pela camada de transporte. Ele geralmente é um número seqüencial randômico acima de 1024;
3. O endereço IP de origem é atribuído pela camada 3.

O UDP (Protocolo de Datagrama de Usuário) é um protocolo sem conexão não confiável para aplicações que não necessitam nem de controle de fluxo nem da manutenção da seqüência das mensagens enviadas. Ele é amplamente usado em aplicações em que a entrega imediata é mais importante do que a entrega precisa, como a transmissão de dados de voz ou vídeo. O UDP foi definido na RFC 768.

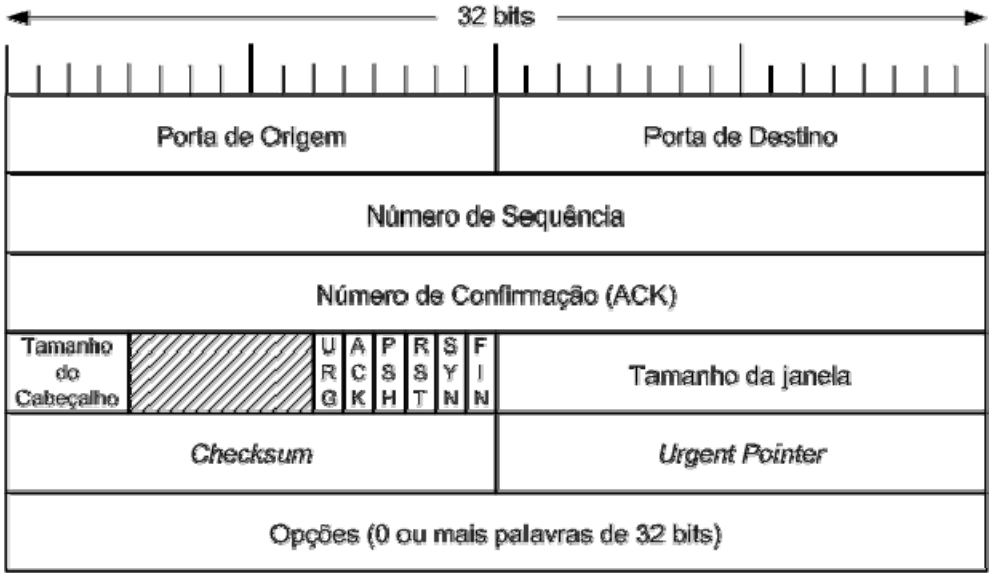
O protocolo UDP é muito mais simples que o TCP, isto se deve ao fato dele não necessitar do estabelecimento de uma conexão (sinalização), controle de fluxo, controle de erros, retransmissão e sequênciamento dos dados. Todas essas funcionalidades são deixadas a cargo da aplicação desenvolvida. Devido a esta simplicidade seu cabeçalho possui apenas 8 bytes, composto por:



- Porta de Origem e Porta de Destino – Indicam os pares de porta que estão executando a comunicação;
- Comprimento – Indica o comprimento de todo o datagrama isto é, cabeçalho e dados;
- Checksum – Verificação de integridade do datagrama;

O TCP (Protocolo de Controle de Transmissão) foi projetado para oferecer um fluxo de bytes fim a fim confiável em uma inter-rede não confiável. Ele é um protocolo orientado a conexão que permite a entrega sem erros de um fluxo de bytes originado de uma determinada máquina para qualquer computados da rede. Esse protocolo fragmenta o fluxo de entrada em mensagens e passa cada uma delas para a camada de redes. No destino, o processo TCP remonta as mensagens recebidas gerando o fluxo de saída. O TCP foi projetado para se adaptar dinamicamente às propriedades da camada de rede e ser robusto diante dos muitos tipos de falhas que podem ocorrer.

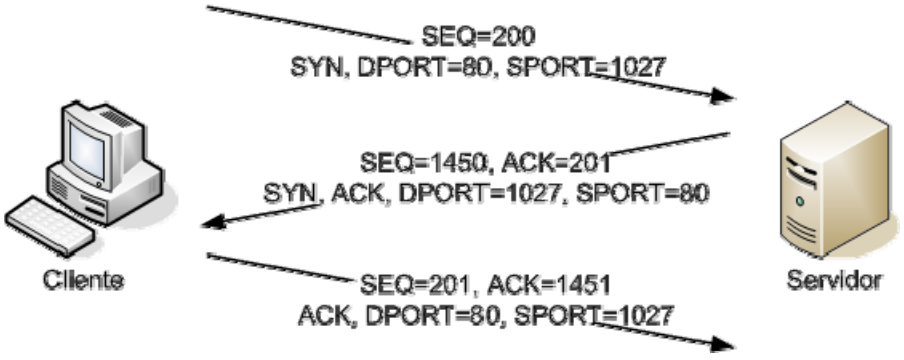
O TCP foi formalmente definido na RFC 793, posteriormente alguns erros foram corrigidos e o TCP foi definido na RFC 1122. O protocolo TCP define um cabeçalho para suas mensagens composto dos seguintes campos:



Cabeçalho TCP

- **Porta de Origem e Porta de Destino** – identifica os pontos terminais locais da conexão;
- **Número de Seqüência –** Identifica o fragmento dentro de todo o fluxo gerado;
- **Numero de Confirmação –** Indica qual o próximo byte esperado;
- **Tamanho do Cabeçalho** – Informa quantas palavras de 32 bits compõem o cabeçalho TCP;
- **URG** – Indica a utilização do *urgent pointer*;
- **ACK** – É utilizado para indicar que este segmento é um ACK e que o campo Número de Confirmação deve ser interpretado;
- **PSH** – Indica que este segmento não deve ser enfileirado como todos os outros, mas sim posto à frente na fila;
- **RST** – É utilizado para reiniciar uma conexão que tenha ficado confusa devido a falhas no *host* ou por qualquer outra razão;
- **SYN** – Este bit é utilizado para indicar um pedido de conexão e a confirmação da conexão;
- **FIN –** Utilizado para indicar que o emissor não possui mais dados para enviar e deseja finalizar a conexão;
- **Tamanho da Janela** – Indica quantos bytes podem ser enviados a partir do byte confirmado. Este campo é utilizado no controle de fluxo do TCP;
- ***Checksum*** – Indicador de integridade do segmento;
- ***Urgent Pointer*** – Indica um deslocamento de bytes a partir do número de seqüência atual em que os dados urgentes devem ser encontrados;
- **Opções** – Projetado para que o TCP possa oferecer recursos extras que não foram previstos em seu protocolo;
-

O estabelecimento de uma conexão TCP ocorre antes que qualquer outro recurso TCP possa começar seu trabalho. O estabelecimento da conexão se fundamenta no processo de inicialização dos campos referentes à seqüência, aos ACKs e na troca dos números de *sockets* usados. As conexões são estabelecidas no TCP por meio do *three way handshake* (*handshake* de três vias).



O estabelecimento da conexão é feito usando dois bits na cabeçalho TCP: SYN e ACK. Um segmento que possua a *flag* SYN ativa sinaliza uma requisição de sincronia do número de seqüência. Essa sincronização é necessária em ambos os sentidos, pois origem e destino utilizam números de seqüência distintos. Cada pedido de conexão é seguido de uma confirmação utilizando o bit ACK. O segundo segmento do *three way handshake* exerce

as duas funções ao mesmo tempo: Confirma a sincronização do servidor com o cliente e requisita a sincronização do cliente com o servidor.

É interessante aqui analisar a imagem acima. Basicamente o Three way handshake 'simula' um acordo. O Cliente pergunta pro servidor: "Você está ai?" e servidor responde: "Sim estou...". Depois o servidor pergunta: "Você está ai?" e o cliente responde: "Sim estou...". Mas fica a dúvida: Como tem 4 sentenças em apenas 3 trocas de mensagens? Simples, a segunda mensagem contém uma resposta e uma pergunta. Como podemos verificar isso? Através das Flags:

Cliente: Servidor, você está ai?? (SYN)
Servidor: Sim estou... (ACK) E você, está ai? (SYN)
Cliente: Sim, estou... (ACK)

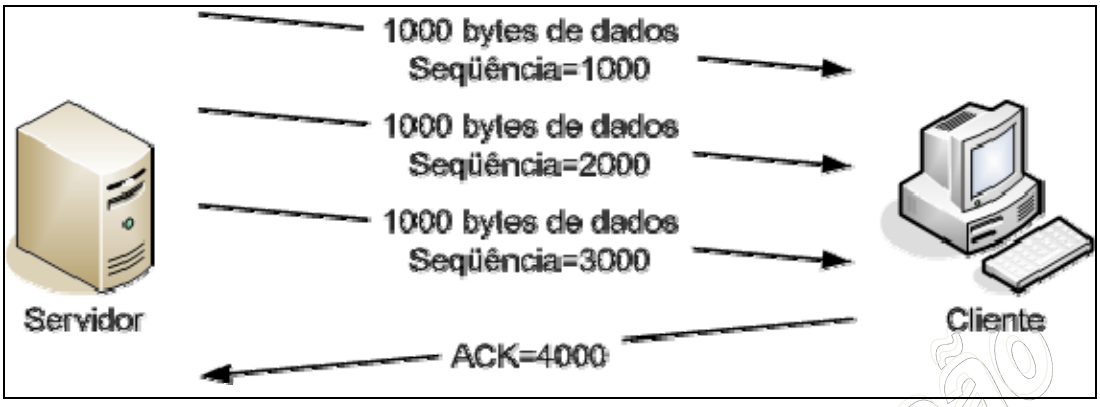
Mas porque o servidor precisa perguntar se o cliente está lá? Pela simples necessidade de sincronização do número de seqüência. O número de seqüência é utilizado para garantir a entrega de todas as mensagens. Vamos imaginar da seguinte forma:

Cliente: Cambio servidor, mensagem 200 (Numero de seqüência do cliente), o senhor está disponível (SYN)?
Servidor: Positivo cliente! Mensagem 1450 (Numero de seqüência do servidor) Prossiga com a mensagem 2001 (ACK=201), cambio.
Cliente: Positivo servidor! Mensagem 201 (numero de sequencia), confirmando número da próxima mensagem: 1451, cambio!

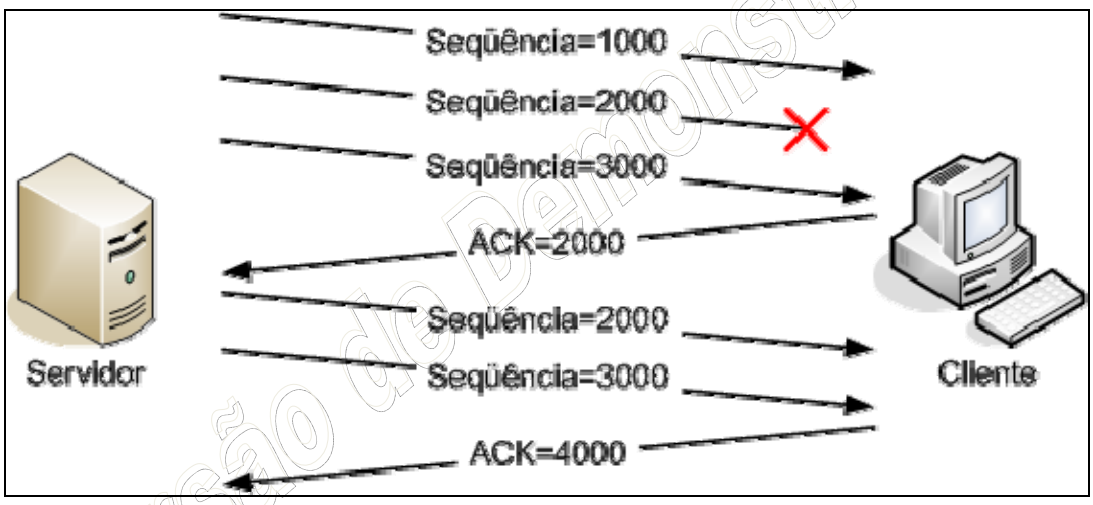
Dessa forma eles trocam o número de seqüência, que tem como função "enumerar" as mensagens de cada um. Por exemplo, se a última mensagem foi a 201 e a mensagem que chegou pro servidor foi a 203, ele tem completa certeza que uma mensagem (202) se perdeu no caminho! Então basta somente solicitar uma retransmissão. O número de seqüência nem sempre é incrementado por 1, ele pode ser incrementado com base no número de bytes enviados pela origem.

O ACK tem como objetivo solicitar a continuidade das mensagens. Podemos interpretar um ACK=210 como sendo: "Pronto, recebi até a 209, pode mandar a 210". Isso vai ser demonstrado com mais calma para frente.

O TCP também proporciona uma transferência confiável de dados, o que também é chamado de confiabilidade ou recuperação de erros. Para conseguir a confiabilidade o TCP enumera os bytes de dados usando os campos referentes à seqüência e aos ACKs no cabeçalho TCP. O TCP alcança a confiabilidade em ambas as direções, usando um campo referente ao número de seqüência de uma direção, combinado com o campo referente ao ACK na direção oposta.

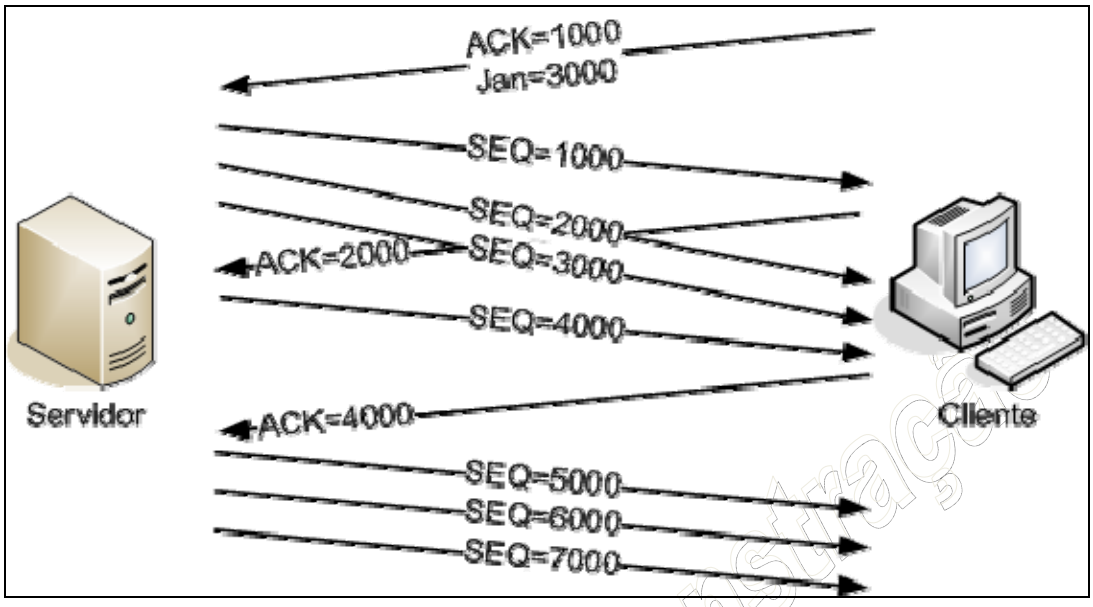


Como dito anteriormente o campo ACK implica o próximo byte a ser recebido. O número de seqüência indica o número do primeiro byte do segmento correspondente à sua posição no fluxo de dados



O TCP implementa o controle de fluxo utilizando os dados dos campos Número de Seqüência, Número de Confirmação e Tamanha da Janela. O controle de fluxo no TCP pode utilizar uma janela de tamanho fixo ou uma janela deslizante.

O campo Tamanho da Janela indica, em tempo real, o número máximo de bytes sem confirmação que podem ser enviados. Com a utilização de janelas um emissor só poderá enviar o número de bytes, previsto na janela, antes de receber alguma confirmação.



Comunicação com janela fixa

Podemos interpretar a imagem a cima da seguinte forma: O Cliente fala pro servidor: Olha servidor, to meio ocupado mas quero continuar esse donwload. Então não me envia mais que 3000 (tamanho da janela) bytes não confirmados OK?

O servidor envia 3 pacotes cada um com 1000 bytes, mas por algum motivo o primeiro chega mais rápido e os dois últimos demoraram um pouco.

O cliente recebe o pacote de seqüência 1000 espera e não recebe mais nada. Então ele envia uma confirmação: “Bora ai cara!! Eu falei 3000!! Só recebi 1000, manda o próximo (ACK)”

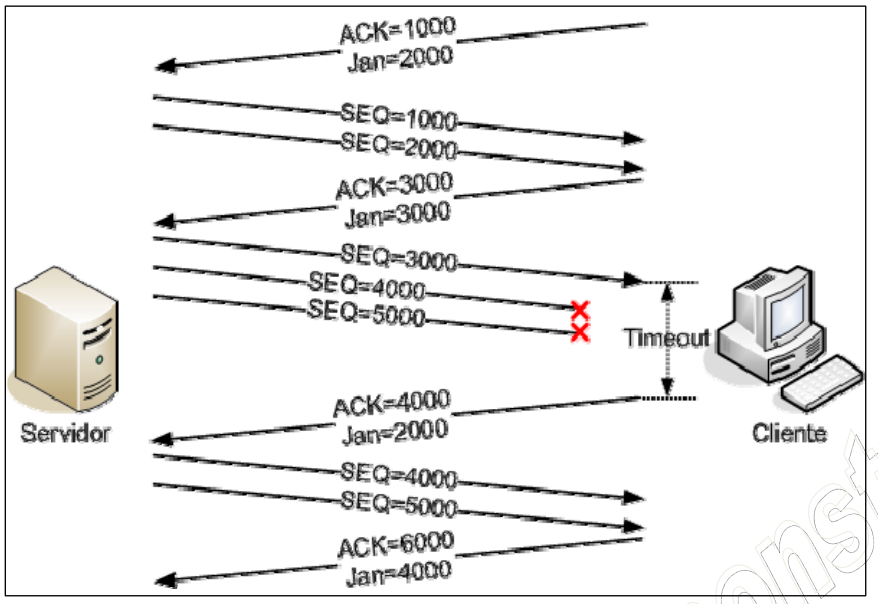
O servidor recebe esse ACK e verifica que enviou 3000 mas só 1000 foram confirmados, ou seja tem 2000 não confirmados e 1000 de espaço livre. Então ele envia mais 1000.

O cliente de repente recebe todos os pacotes, então ele responde: “Beleza, recebi até o 4000, manda o 5000!!! (ACK)”

Então o servidor manda mais 3000 bytes.

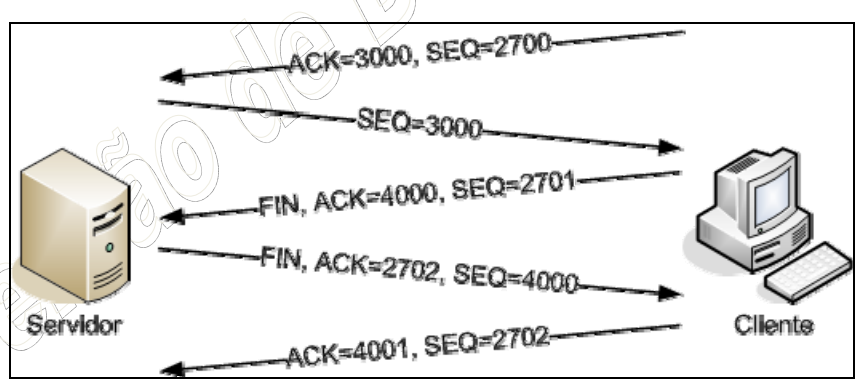
Caso o protocolo TCP esteja utilizando janelas deslizantes o tamanho da janela irá variar ao longo de uma transmissão. Ao iniciar uma conexão a janela começa pequena e aumentará gradativamente até que ocorram erros ou o destinatário seja sobrecarregado. Ao serem detectados erros a janela diminui, após um tempo o tamanho da janela começa a aumentar novamente. Caso o destinatário perceba uma sobrecarga, no próximo ACK enviado por ele haverá um novo tamanho de janela, o qual ele acredita ser apropriado para sua recuperação. Caso seja enviado um valor igual à zero o destinatário esta informando que não possui condições de processar mais dados e a comunicação estará suspensa até que o remetente receba um tamanho de janela diferente de zero.

Aqui vemos um inicio com uma janela de 2000, logo depois sendo incrementada para 3000. Como houve um timeout, o cliente imagina que pode ter havido algum problema e solicita a redução do tamanho da janela para 2000. Como ele percebeu que tudo ocorreu bem, ele solicita uma aumento para 4000.

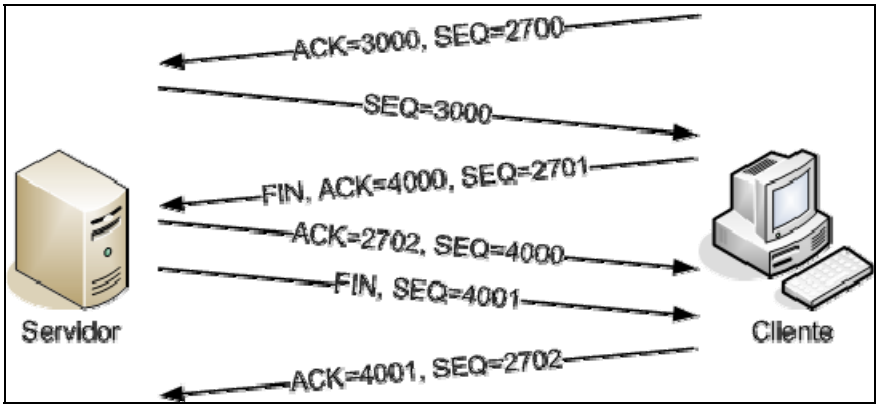


É por causa desse comportamento que o tempo de download nunca é confiável, pois o número de bytes transmitidos é variável. E também por isso que no início o download começa com uma taxa de transferência baixa e vai aumentando aos poucos

A finalização de uma conexão TCP é feita por meio de uma confirmação de três ou quatro vias. Nela é utilizada a *flag* FIN para indicar um pedido de desconexão. Este procedimento deve ser feito em ambas as direções.



Finalização em três vias



Finalização em quatro vias

9.7 NÍVEL DE APLICAÇÃO

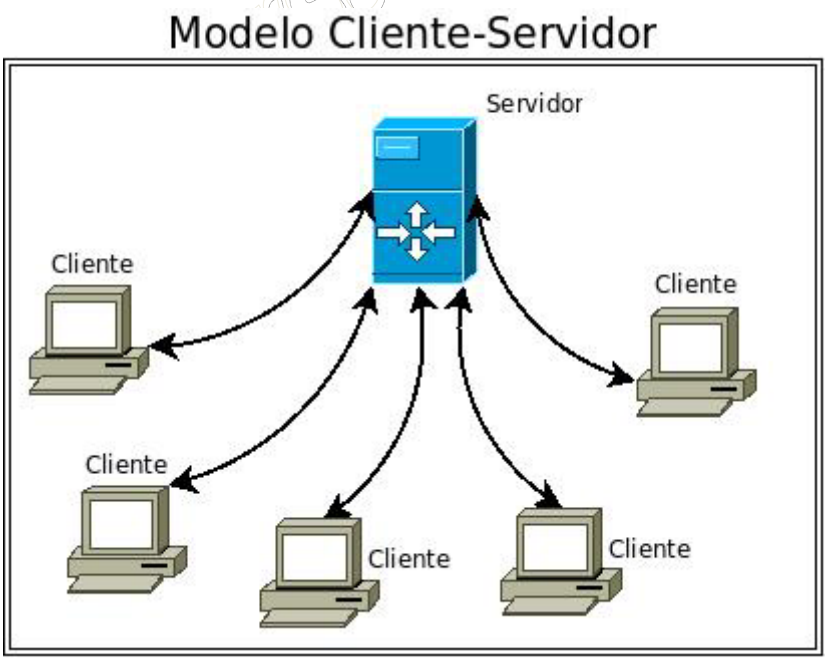
Uma Arquitetura de Aplicação define a estrutura de comunicação entre os utilizadores da aplicação. Existem basicamente três tipos de arquitetura: Cliente-Servidor, Peer-to-Peer e uma arquitetura híbrida, que é uma mescla das outros duas. Ao contrario de uma arquitetura de rede, que é fixa, ou seja, provê um conjunto específico de serviços as aplicaçõoes, a arquitetura de aplicação deve ser escolhida pelo desenvolvedor da aplicação, determinando o modo que a aplicação vai se comportar nos sistemas finais em uma rede.

Com essa classificação segundo a arquitetura (cliente-servidor, P2P ou híbrida) pode-se entender melhor como se comportam as aplicações em uma rede. Em qualquer uma dessas arquiteturas, uma aplicação se comunica através de pares de processos, onde um é rotulado cliente e outro servidor. Mesmo em uma aplicação do tipo P2P, o par que solicita um arquivo de outra máquina, é denominado cliente, e o outro que fornece é o servidor.

O modelo cliente-servidor praticamente ocupava a única possibilidade e acabava assumindo como unanimidade o posto de arquitetura de aplicação, isso ocorria devido a computadores poderosos, com muita memória, serem muito caros. Com isso, a tendência era que existissem computadores potentes que centralizassem esses efeitos, por isso MainFrames eram utilizados para armazenar dados de clientes para fazer operações remotas.

Na atualidade, apesar do avanço da tecnologia, trazendo computadores pessoais com maior possibilidade de processamento e de memória, com custo baixo, esse modelo ainda se apresenta com muita força e aparentemente terá forças para continuar por muito tempo ainda.

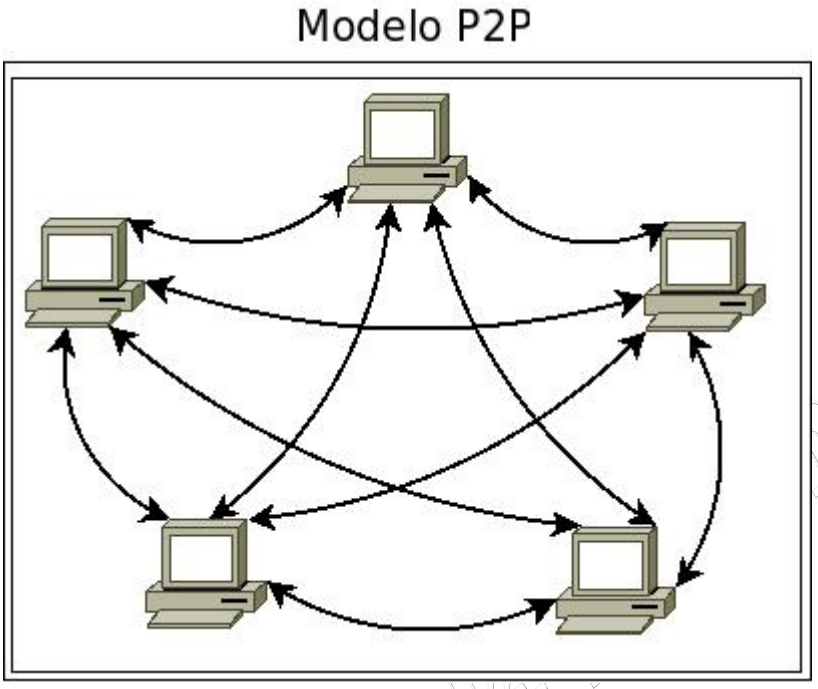
No modelo de arquitetura Cliente-Servidor, existem dois processos envolvidos, um no host cliente e um outro no host servidor. A comunicação acontece quando um cliente envia uma solicitação pela rede ao processo servidor, e então o processo servidor recebe a mensagem, e executa o trabalho solicitado ou procura pelos dados requisitados e envia uma resposta de volta ao cliente, que estava aguardando. Nesta arquitetura o servidor tem uma aplicação que fornece um determinado serviço e os clientes tem aplicações que utilizam este serviço. Uma característica desta arquitetura, é que um cliente não se comunica com outro cliente, e o servidor, que tem um endereço fixo, esta sempre em funcionamento.



Quase sempre um único servidor é incapaz de suportar as requisições de todos os clientes, devido a isso, na maioria dos casos são utilizados vários servidores que constituem um servidor virtual (server farm). Um exemplo claro de aplicação Cliente-Sevidor é a comunicação entre um browser, que é usado para visualizar

páginas da internet, em um servidor web. Neste tipo de aplicação o cliente (browser) e o servidor (servidor web) comunicam-se trocando mensagens através do protocolo HTTP.

A arquitetura P2P (Peer-to-Peer) consiste em uma comunicação direta entre os clientes, não existe nenhuma divisão fixa entre cliente e servidor. Cada par (peer) ativo requisita e fornece dados a rede, desta forma não existe a dependência do servidor, isso aumenta significativamente a largura de banda e a redução de recursos. Esse tipo de arquitetura é utilizado principalmente por aplicações de compartilhamento de conteúdo, como arquivos contendo áudio, vídeo, dados ou qualquer coisa em formato digital. Outras aplicações orientadas a comunicações de dados, como a telefonia digital, videotelefonia e rádio pela internet também utilizam esta arquitetura. Como exemplo podemos citar o protocolo BitTorrent que utiliza a arquitetura peer-to-peer para compartilhamento de grandes quantidades de dados. Neste exemplo um cliente é capaz de preparar e transmitir qualquer tipo de ficheiro de dados através de uma rede, utilizando o protocolo BitTorrent.

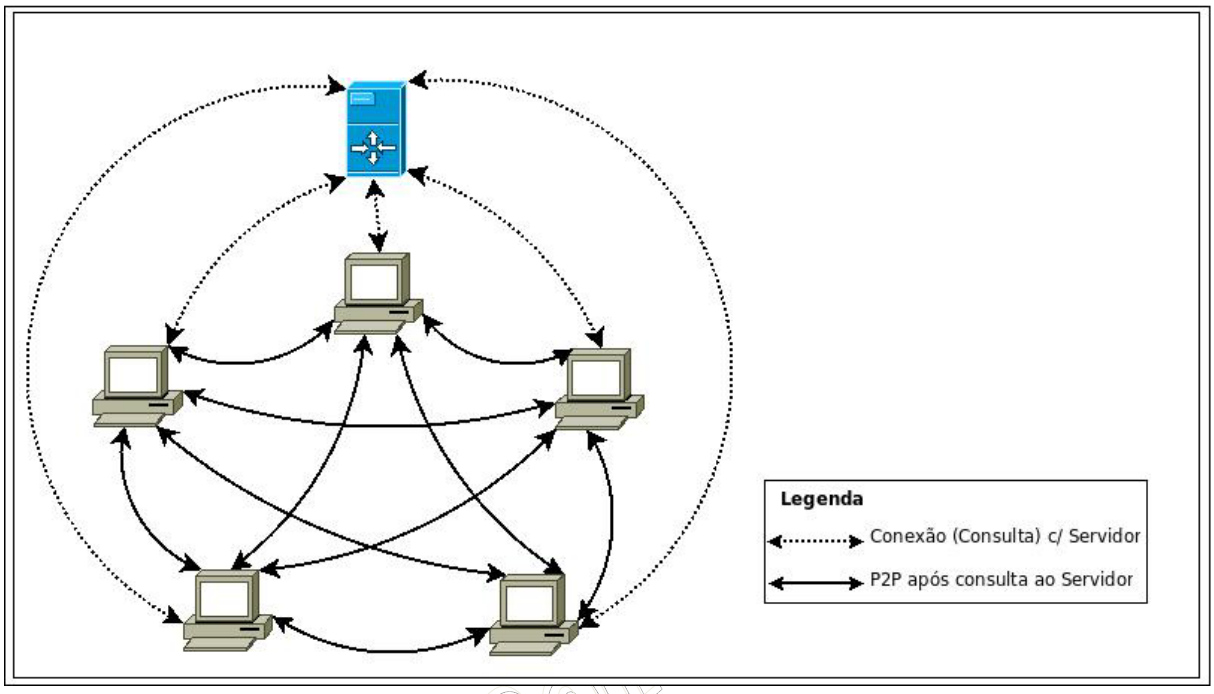


Um peer (par) é qualquer computador que esteja executando uma instância de um cliente. Para compartilhar um arquivo ou grupo de arquivos, um nó primeiro cria um pequeno arquivo chamado "torrent" (por exemplo, Meuarquivo.torrent). Este arquivo contém metadados sobre os arquivos a serem compartilhados e sobre o tracker, que é o computador que coordena a distribuição dos arquivos. As pessoas que querem fazer o download do arquivo devem primeiro obter o arquivo torrent, e depois se conectar ao tracker, que lhes diz a partir de quais outros pares que se pode baixar os pedaços do arquivo.

Com uma pesquisa realizada pela empresa Xerox, foi detectado que pelo menos 70% dos usuários de P2P não compartilhavam arquivo, enquanto apenas 1% compartilhavam 50% destes, ou seja, a teoria que se tinha de “divisão de trabalho” pelos clientes, não valia na prática. Para isso então, buscou-se uma solução, e esta solução, representou a utilização da arquitetura do tipo híbrida.

Uma híbrida, mescla das outras duas: cliente-servidor/P2P. Esta arquitetura utiliza, por exemplo, para transferência de arquivos o P2P e a arquitetura cliente/servidor para pesquisar quais peers contêm o arquivo desejado. Uma aplicação muito utilizada neste tipo de arquitetura é a de mensagem instantânea. O Windows Live Messenger e o aMSN são bons exemplos, onde usuários podem bater papo online instantaneamente em tempo real. A comunicação desta aplicação é tipicamente P2P, no entanto, para iniciar uma comunicação, um usuário registra-se em um servidor, e verifica quem da sua lista de contatos também está registrado, para a partir de então começar uma comunicação. Essas aplicações também disponibilizam transferência de arquivos, suporte a grupos, emoticons, histórico de chat, suporte a conferência, suporte a Proxy, e outras ferramentas.

Modelo Híbrido



Na Internet, as aplicações devem "conversar" entre si, ou seja, o que o usuário deseja deve ser entendido pela outra máquina e respondido. Essa comunicação é feita entre os processos, através da troca de mensagens. O remetente cria mensagens com seus pedidos ao destinatário, que recebe e gera as suas mensagens para responder (ou não) a solicitação.

Por exemplo, numa comunicação Web, o cliente solicita uma página da Internet, através de um determinado tipo de mensagem (no caso, uma requisição HTTP). O servidor recebe a requisição, e envia uma mensagem com a página para o cliente (através de uma resposta HTTP). Porém, se ocorre um erro, o servidor envia mensagens dizendo ao cliente que ouve algum erro.

Geralmente, a comunicação consiste em pares de processos, onde um processo em cada lado envia mensagens para o outro. Isso ocorre na rede através dos sockets, que são os "porta-vozes" de cada host para uma determinada aplicação.

Para que haja essa comunicação, é necessário que os hosts se identifiquem. Para isso, usam o endereço IP. Porém, é necessário também identificar qual processo naquela máquina irá levar as mensagens à aplicação, e essa identificação é chamada de número (ou endereço) de porta.

Para que dois processos se comuniquem, eles devem trocar mensagens. Porém, é necessário haver regras que padronizem como serão trocadas e tratadas essas mensagens. Por isso, existem os protocolos da camada de aplicação. Como em Tanenbaum[2], "mesmo na camada de aplicação existe a necessidade de protocolos de suporte, a fim de permitir que as aplicações funcionem." É necessário definir os tipos de mensagens a serem trocadas, a sintaxe dos vários tipos de mensagens, a semântica dos campos que compõem as mensagens e as regras que determinam quando e como um processo envia e responde as mensagens. No entanto, como explica Kurose, é importante não confundir os protocolos de camada de aplicação com as aplicações. São conceitos diferentes, apesar de os protocolos serem uma parcela significativa de uma aplicação. Uma aplicação é a interface com o usuário, ou seja, aquilo que é realmente acessado. Os protocolos se responsabilizam por definir como os processos irão se comunicar e como irão

tratar as mensagens, para expor o que foi solicitado pelo usuário em sua aplicação. Por exemplo, para acessar uma página Web, um usuário executa um programa Browser e solicita uma página. O Browser usa o protocolo HTTP para enviar o pedido da página, assim como o servidor usa o mesmo protocolo para aceitar a requisição e devolver a página solicitada. O Browser interpreta a mensagem vinda do servidor e apresenta a página.

Dentre os protocolos de aplicação, pode-se citar: HTTP (HyperText Transfer Protocol), HTTPS (HyperText Transfer Protocol over Secure Socket Layer), FTP (File Transfer Protocol), SMTP (Simple Mail Transfer Protocol), Telnet, POP3 (Post Office Protocol version 3), e muitos outros. Vejamos a seguir alguns dos protocolos mais importantes de aplicação.

9.7.1 NCP

Network Control Protocol, primeiro protocolo servidor a servidor da ARPANET. Ele foi criado em dezembro de 1971, pelo *Network Working Group* (NWG). Não está mais em atividade.

9.7.2 Telnet e SSH

TELNET (Protocolo de Terminal Virtual) é o protocolo Internet 23/TCP para estabelecer a conexão entre computadores. Através dessa conexão remota, podem-se executar programas e comandos em outra máquina, como se o teclado de seu computador estivesse ligado diretamente a ela.

O visual de uma conexão via Telnet é semelhante ao que se tem em BBS's de interface MSDOS, e a operação do computador remoto se dá da mesma forma, ou seja, através de uma linha de comandos Unix ou a partir de um menu de comandos disponíveis que sempre se apresenta em algum lugar da tela (esta última forma é a mais comum em servidores que permitem acesso público).

O Telnet pode ser usado para a pesquisa de informações e transferência de arquivos - tudo depende do que o computador ao qual você está conectado permitir que você faça. Ele também é muito usado por operadores de sistemas (Sysop's) a fim de fazer algum tipo de manutenção (se você pensa que o Sysop de seu provedor sai de casa toda vez que tem algum problema nos servidores, está muito enganado; muitas vezes ele faz a manutenção de casa mesmo, via Telnet !)

Telnet é um protocolo pouco seguro, por isso quase se deixou de usar. Agora o típico é utilizar SSH (Secure Shell) que é outro protocolo muito similar, embora adotando mecanismos de proteção baseado em criptografia.

Quando nos conectamos por telnet com um servidor temos que enviar nosso nome de usuário e senha. Estes dados críticos se enviam por meio de texto claro, sem nenhum tipo de encriptação, por isso qualquer um poderia lê-los se estiver "escutando" nossas comunicações.

Para evitar este grave problema de segurança se utiliza SSH, que é um protocolo de comunicação em redes muito parecido, porém onde todas as comunicações viajam de maneira encriptada.

SSH costuma trabalhar na porta 22/TCP e os programas que permitem fazer telnet, o mais normal é que permitam também fazer SSH. Por exemplo, o software Putty realiza tanto Telnet quanto SSH, o que às vezes se chama "telnet por ssh".

As características do SSH são: cifragem dos dados utilizando algoritmos livres de patente (DES, 3DES, AES, entre outros; X11 forwarding, protege o display das conexões X11 em Unix/Linux; Port Forwarding, permite encapsular a comunicação de outros protocolos, como o FTP, POP3, SMTP, utilizando TCP através de um canal cifrado; Autenticação forte, protege contra falsificação do endereço de origem; Agent Forwarding, permite o uso de certificados digitais para autenticação de clientes móveis; Compressão de dados.

9.7.3 DNS e DNSSec

Quando você visita um site através do seu navegador ou quando envia um email, a internet precisa saber em qual servidor o site e o e-mail estão armazenados para poder responder à sua solicitação. A informação da localização destes servidores está em um servidor chamado DNS (Domain Name Server).

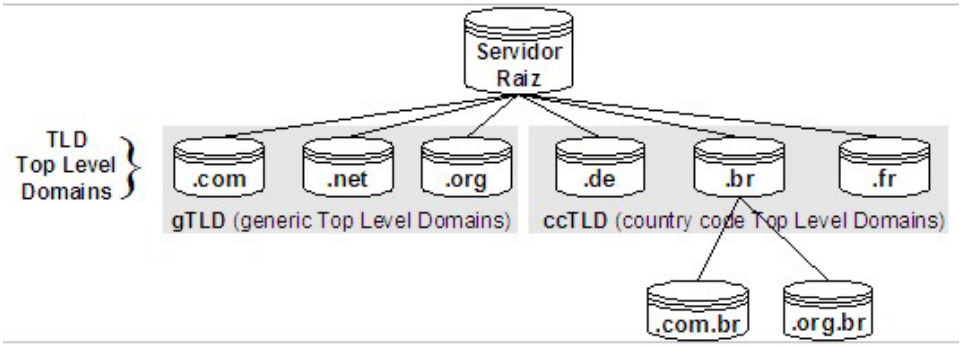
Cada domínio possui um registro no DNS que define qual o endereço IP do servidor de hospedagem e o IP do servidor de e-mail que responderão por este domínio. O processo para a descoberta dos servidores que respondem por um domínio é denominado “resolução do nome” ou “resolução do domínio”.

Os navegadores e os sistemas clientes de e-mail solicitam que a internet faça a resolução do domínio para apresentar um site, ou enviar um e-mail. Esse processo é totalmente transparente para o usuário, que apenas digita o site que quer visitar e o navegador descobre em qual servidor o site está hospedado e em seguida solicita para o servidor de hospedagem que envie a página inicial.

Por segurança, um domínio pode definir vários servidores DNS. O DNS primário é o primeiro sistema a ser consultado no momento da resolução do nome, caso o servidor DNS primário esteja em manutenção, o servidor DNS secundário é consultado, e assim sucessivamente.

Devido ao intenso tráfego da internet e devido à segurança da rede, a estrutura do banco de dados DNS é distribuída e hierárquica. Ou seja, ao invés de um banco de dados central e único com informações de todos os domínios, a resolução ocorre consultando-se diversos servidores DNS e sua resolução é hierárquica (um servidor DNS pode apontar para outro servidor DNS e assim sucessivamente).

A estrutura hierárquica equivale a uma árvore invertida, ou seja, existe um servidor principal que aponta para um secundário que aponta para um terceiro e assim sucessivamente. O servidor DNS que está no topo da internet é o servidor raiz.



O servidor raiz da internet possui uma tabela que indica qual DNS será responsável pela resolução dos domínios para cada extensão de domínio (Top Level Domain) diferente.

A tabela em si é muito pequena, possui apenas uma entrada para cada Top Level Domain existente. Os Top Level Domains são de dois tipos: gTLDs (Generic Top Level Domains - domínios genéricos usados

no mundo todo) e ccTLDs (Country Code Top Level Domains - extensões de domínios administrados pelos países).

Por exemplo: todos os domínios terminados em .com serão respondidos pelos servidores da VeriSign; os domínios .br serão respondidos pelos servidores do Registro.br e assim sucessivamente. Cada gTLD ou ccTLD tem apenas uma entrada neste banco de dados.

Por segurança, o servidor raiz foi replicado em 13 servidores raízes diferentes espalhados pelo mundo e duas vezes ao dia seu conteúdo é automaticamente replicado.

Foi convencionado que cada servidor raiz seria chamado por uma letra do alfabeto (Servidor A, Servidor B etc…). Mesmo um determinado servidor raiz, o servidor raiz A, por exemplo, pode ser replicado em várias regiões do mundo, para assegurar que o tempo para a resolução de um domínio seja rápido (baixa latência).

Bem, então na verdade existem treze servidores raiz principais e dezenas de cópias espalhadas pelo mundo. Veja na imagem abaixo a plotagem dos servidores raízes e suas cópias em funcionamento no mundo.

Os grandes provedores de acesso e empresas de telecomunicações arquivam em seus caches (memória temporária) a tabela dos servidores raiz. Portanto, a cada e-mail enviado ou site visitado os servidores raiz não são obrigatoriamente consultados. Na verdade, o volume de consultas a estes servidores é muito pequeno, já que sua tabela é alterada apenas quando um novo top level domain é criado. Quem realmente processa o maior volume de queries para resolução de nomes são os servidores dos TLDs (Top Level Domains).

Por exemplo: um servidor raiz normalmente recebe 500 queries por dia e os servidores da VeriSign (responsável pela resolução dos domínios .com) recebem bilhões de queries diariamente.

A estrutura hierárquica de resolução de nomes, onde um DNS aponta para outro DNS, possui um problema intrínseco de segurança. Imagine a hipótese que um provedor de acesso capture uma querie para resolução de um nome e inadvertidamente responda com um endereço errado de onde o site esteja hospedado. Neste exemplo, você poderia solicitar no seu navegador o endereço www.itau.com.br e o provedor fornecer por erro www.brasdeco.com.br, ou pior, um site phishing, que simula o site do banco Itaú.

Um dos maiores problemas desta hipótese é que realmente seria impossível identificar que o provedor de acesso fez isso. Portanto, para dar segurança a estrutura de resolução de nomes a IETF (Internet Engineering Task Force) criou uma extensão do uso atual do DNS denominado DNSSEC.

A extensão DNSSEC autentica as informações do DNS e garante que estas informações são autênticas e íntegras. Sua adoção depende de cada Top Level Domain. O Registro.br, responsável pela

administração dos domínios .br já começou a permitir o registro de domínios com o DNSSEC para algumas extensões como .blog.br, .eng.br etc.

O mercado aguarda a liberação do uso do DNSSEC para a extensão .com.br, de longe a mais utilizada no país. O mercado bancário e financeiro devem ser os primeiros a aderir ao DNSSEC e devem solicitar para que as empresas responsáveis pela sua hospedagem façam esta implementação extra de segurança.

9.7.4 UUCP e SMTP

O primeiro protocolo de transferencia desenvolvido foi o UUCP (Unix to Unix CoPy), sob regência do RFC 976. Surgiu e foi bastante difundido por volta dos anos 80.

Inicialmente foi utilizado na ARPANET, para troca de mensagens entre Universidades. Como funcionava sobre redes comutadas por circuitos (e portanto a tarifação era por tempo de conexão), e ainda por ser necessário uma conexão entre cada cliente, que muitas vezes estavam em outros países, era comum implantar um sistema concentrador de atividades.

Este concentrador sincronizava-se com os clientes e armazenava as funções pedidas, como envio de e-mails e transferencia de arquivos, e em determinada hora conectava-se e realizava as funções da fila. Após concluído, desconectava-se e voltava a armazenar as funções.

Este comportamento conferia uma certa desvantagem por não ser em tempo real, com atrasos de várias horas, mas com certeza havia grande vantagem sobre os correios convencionais, que demoravam dias ou meses. Os e-mails conforme esta tecnologia eram formados pelo nome da máquina seguido de exclamação e do nome do usuário (Exemplo: dominio.com.br!nome.de.usuario. Neste tipo de protocolo, era extremamente comum o uso de servidores intermediários, o que barateava a comunicação. Em geral, um servidor só possuía acesso aos seus adjacentes. Se eu fosse mandar um e-mail para a China por exemplo, deveria utilizar o endereço de destinatário ServidorBrasil!ServidorEuropa!ServidorLesteEuropa!ServidorChina!usuário. Esta prática aumentava ainda mais o atraso com que as mensagens chegavam.

Como esta definição de rotas estáticas era bastante trabalhosa, começaram a ser implantados na rede hops, que eram máquinas capazes de interpretar as rotas e reescrever outras mais rápidas e menos congestionadas, o que melhorou a velocidade da comunicação e reduziu custos.

Atualmente este protocolo ainda é utilizado em redes corporativas e alguns sistemas devido ao baixo custo, gerenciamento não-persistente de filas, porém com adaptações para uso sobre o protocolo TCP/IP. Gradativamente, no entanto, ela vem sido substituída por técnicas mais modernas.

A tecnologia utilizada pela NASA para comunicação com suas sondas e satélites é similar à UUCP.

SMTP, ou simple mail transfer protocol, conforme define o RFC 2821, é o protocolo mais utilizado atualmente para transmissão de mensagens de correio eletrônico.

O protocolo é utilizado pelo MTA para transferir a mensagem, e ele serve justamente para definir padrões de como entregar, e como interpretar os ados enviados. O padrão exige a codificação de binário em ASCII, e decodificação ASCII para binário na passagem ao MDA.

Em geral, uma transferencia SMTP é direta entre o servidor de origem e o de destino, não passando por nenhum intermediário. Os servidores armazenam as mensagens caso não possam ser entregues de imediato, por qualquer falha ou impedimento. A conexão é feita na porta TCP 25.

A comunicação entre servidores SMTP é estabelecida sobre o protocolo TCP/IP, com a identificação dos conectantes. Após estabelecida a conexão, há a troca de comandos entre o cliente e o servidor, iniciando-se com a identificação do remetente, após do destinatário, e por fim a mensagem.

Por se tratar de uma conexão persistente, podem ser enviadas diversas mensagens sequencialmente, bastando apenas especificar o remetente, destinatário e mensagem dos demais emails antes do comando de encerrar a conexão (quit).

S: 220 www.example.com ESMTP Postfix
C: HELO mydomain.com
S: 250 Hello mydomain.com
C: MAIL FROM: sender@mydomain.com
S: 250 Ok
C: RCPT TO: friend@example.com
S: 250 Ok
C: DATA
S: 354 End data with <CR><LF>.<CR><LF>
C: Subject: test message
C: From: sender@mydomain.com
C: To: friend@example.com
C:
C: Hello,
C: This is a test.
C: Goodbye.
C: .
S: 250 Ok: queued as 12345
C: quit
S: 221 Bye

9.7.5 FTP

FTP significa *File Transfer Protocol* (Protocolo de Transferência de Arquivos), e é uma forma bastante rápida e versátil de transferir arquivos (também conhecidos como ficheiros), sendo uma das mais usadas na internet.

Pode referir-se tanto ao protocolo quanto ao programa que implementa este protocolo (Servidor FTP, neste caso, tradicionalmente aparece em letras minúsculas, por influência do programa de transferência de arquivos do Unix).

A transferência de dados em redes de computadores envolve normalmente transferência de arquivos e acesso a sistemas de arquivos remotos (com a mesma interface usada nos arquivos locais). O FTP (RFC 959) é baseado no TCP, mas é anterior à pilha de protocolos TCP/IP, sendo posteriormente adaptado para o TCP/IP. É o padrão da pilha TCP/IP para transferir arquivos, é um protocolo genérico independente de hardware e do sistema operacional e transfere arquivos por livre arbítrio, tendo em conta restrições de acesso e propriedades dos mesmos.

A transferência de arquivos dá-se entre um computador chamado "cliente" (aquele que solicita a conexão para a transferência de dados) e um servidor (aquele que recebe a solicitação de transferência). O utilizador, através de software específico, pode selecionar quais arquivos enviar ao servidor. Para existir uma conexão ao servidor, o utilizador informa um nome de utilizador (ou username, em inglês) e uma senha (password), bem como o nome correcto do servidor ou seu endereço IP. Se os dados foram informados corretamente, a conexão pode ser estabelecida, utilizando-se um "canal" de comunicação, chamado de porta (port). Tais portas são conexões no qual é possível trocar dados. No caso da comunicação FTP, o padrão para porta é o número 21.

O acesso a servidores FTP pode ocorrer de dois modos: através de uma interface ou através da linha de comando, tanto usuários UNIX como usuários Windows podem acessar através dos dois modos. Embora um pouco complicado, o modo linha de comando está presente em qualquer distribuição UNIX-like e Windows, através do telnet.

A partir de qualquer browser credenciado(Internet Explorer, Firefox, ou mesmo no Windows Explorer) também é possível aceder a um servidor FTP. Basta, para isso, digitar na barra de endereço:

ftp:// [username] : [password] @ [servidor]

O protocolo subjacente ao FTP pode rodar nos modos interativo ou *batch*. O cliente FTP fornece uma interface interativa, enquanto que o MIME e o HTTP usam-no diretamente. O protocolo permite a gravação e obtenção de arquivos, a listagem da pasta e a alteração da pasta de trabalho.

Os servidores de FTP raramente mudam, mas novos clientes FTP aparecem com bastante regularidade. Estes clientes variam no número de comandos que implementam, a maioria dos clientes FTP comerciais implementam apenas um pequeno subgrupo de comandos FTP. Mesmo que o FTP seja um protocolo orientado a linha de comandos, a nova geração dos clientes FTP esconde esta orientação num ambiente gráfico, muitas vezes, muito desenvolvido.

A interface cliente do FTP do BSD UNIX é um padrão por si mesma, possuindo muitos comandos arcaicos: tenex ou carriage control que hoje não têm uso. Os comandos mais usados são o cd, dir, ls, get e put.

O FTP tem particularidades que são hoje pouco comuns. Depois da ativação do ftp, é estabelecida uma conexão ao host remoto. Esta conexão envolve o uso da conta do usuário no host remoto, sendo que alguns servidores FTP disponibilizam *anonymous FTP*.

Certos comandos são os que fazem a transferência bidirecional de arquivos, são eles:

- get do servidor FTP para o host local (mget para mais que um arquivo)
- put para o servidor FTP a partir do host local (mput para mais que um arquivo)

Nota: alguns comandos podem não funcionar com o usuário sendo *anonymous*, pois tal conta tem limitações de direitos a nível do sistema operacional.

A sintaxe dos nomes dos arquivos pode ser incompatível entre diferentes Sistemas Operacionais. O UNIX usa 128 caracteres, maiúsculas e minúsculas, enquanto que o DOS usa 8 + 3 caracteres e apenas maiúsculas. Certos nomes não podem ser usados em alguns sistemas. Devido a isto tudo o BSD ftp define regras para a tradução de nomes.

O FTP permite dois modos de transferência de mensagens FTP: *texto* (com traduções apropriadas) ou *binário* (sem tradução). Cada mensagem do servidor inclui um identificador decimal de 3 dígitos

(exemplo: 226 Transfer complete). Estas mensagens podem ser vistas ou não, usando para isso o modo *verbose* ou *quiet*, respectivamente.

O Servidor remoto aceita uma *conexão de controle* do cliente local. O cliente envia comandos para o servidor e a conexão persiste ao longo de toda a sessão (tratando-se assim de um protocolo que usa o TCP).

O servidor cria uma *conexão de dados* para a transferência de dados, sendo criada uma conexão para cada arquivo transferido. Estes dados são transferidos do servidor para o cliente e vice e versa.

Os comandos estão separados dos dados e o cliente pode enviar comandos durante a transferência de dados. O encerramento da conexão indica o fim do arquivo.

Os comandos abaixo podem ser executados no FTP através da linha de comando. Os comandos do FTP podem ser abreviados, desde que não formem expressões ambiguas.

Podemos ver a seguir um conjunto típico de comandos do FTP obtidos diretamente através do comando *help*.

ftp> help

Os comandos podem estar abreviados. Seguem os comandos:

!	delete	literal	prompt	send
?	debug	ls	put	status
append	dir	mdelete	pwd	trace
ascii	disconnect	mdir	quit	type
bell	get	mget	quote	user
binary	glob	mkdir	recv	verbose
bye	hash	mL	remotehelp	
cd	help	mput	rename	

9.7.6 POP3, IMAP e Webmail

Protocolo de acesso extremamente simples, definido pelo RFC 1939. Seu nome vem da abreviação de Post Office Protocol versão 3.

Há basicamente três passos que devem ser executados: autenticação, transação e atualização. Na autenticação, após estabelecida a conexão, o cliente fornece um nome de usuário e um senha, sem nenhuma obsfuscação. Após, há duas opções para a transação: ler-e-apagar, e ler-e-guardar, o que influencia nos comandos que devem ser passados ao servidor. Na fase de atualização, que ocorre após o término da conexão, o servidor apaga ou marca como lida as mensagens, conforme definido na fase de transação.

Embora seja a 3ª versão deste protocolo, ele é muito simples. Utiliza-se basicamente de 6 comandos: user, pass, list, retr, dele e quit. Responde basicamente de duas formas: err quando um comando está incorreto, e ok quando o comando foi compreendido. A conexão é feita na porta TCP 110. Embora simples, é o mais indicado para pessoas que acessam e-mail de apenas um local.

O protocolo IMAP (Internet Message Access Protocol) é mais robusto que o POP, e está em sua quarta versão, primeira revisão, definido na RFC 3501. Seu poder aumentou sua complexidade relativamente ao POP.

O IMAP é ideal para usuário nômades, que acessam de diversos pontos, pois permite a gerência remota de ações, inclusive entre sessões. Não entendeu? Você organiza sua mensagens na pasta local e elas são organizadas similarmente na sua caixa postal, com comandos do usuário. Há também a vantagem de poder receber somente determinada parte de uma mensagem, nos casos de uma conexão lenta, estreita, ou muito cara (como celular por exemplo). Neste caso o usuário pode filtrar para receber parte da mensagem, escolher quais conteúdos baixar, ou seomente mensagens pequenas.

O poder que este protocolo confere é imenso. Vale a pena ler a RFC 3501.

O webmail ou e-mail sobre HTTP é uma funcionalidade excelente para usuários em trânsito. A transmissão das mensagens para o servidor e da caixa de entrada ao usuário são feitas através do protocolo HTTP, que permite o acesso através de qualquer browser.

Isto confere maior agilidade e portabilidade ao uso do e-mail. É importante lembrar que as trocas entre servidores de webmail continuam sendo feitas através de SMTP.

O webmail pode ser considerada a modalidade de acesso a e-mails mais utilizada atualmente. Muitos webmails utilizam scripts que conferem funcionalidades IMAP ao usuário.

9.7.7 WWW e HTTP

As idéias por trás da Web podem ser identificadas ainda em 1980, na CERN (Suíça), quando Tim Berners-Lee construiu o ENQUIRE. Ainda que diferente da Web atualmente, o projeto continha algumas das mesmas idéias primordiais, e também algumas idéias da web semântica. Seu intento original do sistema foi tornar mais fácil o compartilhamento de documentos de pesquisas entre os colegas.

Em março de 1989, Tim Berners-Lee escreveu uma proposta de gerenciamento de informação, que referenciava o ENQUIRE e descrevia um sistema de informação mais elaborado. Com a ajuda de Robert Cailliau, ele publicou uma proposta mais formal para a World Wide Web (Rede de Alcance Mundial) no final de 1990.

Um computador NeXTcube foi usado por Berners-Lee com primeiro servidor web e também para escrever o primeiro navegador, o WorldWideWeb, em 1990. No final do mesmo ano, Berners-Lee já havia construído todas as ferramentas necessárias para o sistema: o navegador, o servidor e as primeiras páginas web, que descreviam o próprio projeto. Em 6 de agosto de 1991, ele postou um resumo no grupo de notícias alt.hypertext. Essa data marca a estréia da Web como um serviço publicado na Internet.

O conceito crucial do hipertexto originou-se em projetos da década de 1960, como o projeto Xanadu e o NLS. A idéia revolucionária de Tim foi unir o hipertexto e a Internet. Em seu livro *Weaving The Web*, ele explica que sugeriu repetidamente o casamento das tecnologias para membros de ambas as comunidades de desenvolvedores. Como ninguém implementou sua idéia, ele decidiu implementar o projeto por conta própria. No processo, ele desenvolveu um sistema de identificação global e único de recursos, o *Uniform Resource Identifier* (URI).

Sistemas anteriores diferenciavam-se da Web em alguns aspectos. Na Web uma hiperligação é unidirecional enquanto trabalhos anteriores somente tratavam ligações bidirecionais. Isso tornou possível criar uma hiperligação sem qualquer ação do autor do documento sendo ligado, reduzindo significativamente a dificuldade em implementar um servidor Web e um navegador. Por outro lado, o sistema unidirecional é responsável por o que atualmente chama-se hiperligação quebrada, isto é, uma hiperligação que aponta para uma página não disponível devido à evolução contínua dos recursos da Internet com o tempo.

Diferente de sistemas anteriores como o HyperCard, a World Wide Web não era software proprietário, tornando possível a criação de outros sistemas e extensões sem a preocupação de licenciamento. Em 30 de abril de 1993, a CERN anunciou que a World Wide Web seria livre para todos, sem custo. Nos dois meses após o anúncio que o gopher (um dos maiores serviços de troca de informações da época) já não era mais livre, produziu-se uma mudança para a Web. Um antigo navegador popular era o ViolaWWW, que era baseado no HyperCard.

Considera-se que a grande virada da WWW começou com a introdução do Mosaic em 1993, um navegador gráfico desenvolvido por um time de desenvolvedores universitários. Antes de seu lançamento, os gráficos não eram freqüentemente misturados com texto em páginas web.

O HyperText Transfer Protocol é um protocolo de aplicação responsável pelo tratamento de pedidos e respostas entre cliente e servidor na World Wide Web. Ele surgiu da necessidade de distribuir informações pela Internet e para que essa distribuição fosse possível foi necessário criar uma forma padronizada de comunicação entre os clientes e os servidores da Web e entendida por todos os computadores ligados à Internet. Com isso, o protocolo HTTP passou a ser utilizado para a comunicação entre computadores na Internet e a especificar como seriam realizadas as transacções entre clientes e servidores, através do uso de regras básicas.

Este protocolo tem sido usado pela WWW desde 1990. A primeira versão de HTTP, chamada HTTP/0.9, era um protocolo simples para a transferência de dados no formato de texto ASCII pela Internet, através de um único método de requisição, chamado GET. A versão HTTP/1.0 foi desenvolvida entre 1992 e 1996 para suprir a necessidade de transferir não apenas texto. Com essa versão, o protocolo passou a transferir mensagens do tipo MIME (*Multipurpose Internet Mail Extension*) e foram implementados novos métodos de requisição, chamados POST e HEAD.

No HTTP/1.1, versão actual do protocolo descrito na RFC 2616, foi desenvolvido um conjunto de implementações adicionais ao HTTP/1.0, como por exemplo: o uso de conexões persistentes; o uso de servidores *proxy* que permitem uma melhor organização da *cache*; novos métodos de requisições; entre outros. Afirma-se que o HTTP também é usado como um protocolo genérico para comunicação entre os agentes de utilizadores e *proxies*/*gateways* com outros protocolos, como o SMTP, NNTP, FTP, Gopher, e WAIS, permitindo o acesso a recursos disponíveis em aplicações diversas.

O protocolo HTTP faz a comunicação entre o cliente e o servidor através de mensagens. O cliente envia uma mensagem de requisição de um recurso e o servidor envia uma mensagem de resposta ao cliente com a solicitação. Segundo Foscarini, os dois tipos de mensagens existentes no protocolo utilizam um formato genérico, definido na RFC 822, para a transferência de entidades.

Uma mensagem, tanto de requisição quanto de resposta, é composta, conforme definido na RFC 2616, por uma linha inicial, nenhuma ou mais linhas de cabeçalhos, uma linha em branco obrigatória

finalizando o cabeçalho e por fim o corpo da mensagem, opcional em determinados casos. Nessa seção serão apresentados os campos que compõem uma mensagem mais detalhadamente; ou seja, o HTTP apresenta o sítio ou local onde está a página da Internet.

O cabeçalho da mensagem (*header*) é utilizado para transmitir informações adicionais entre o cliente e o servidor. Ele é especificado imediatamente após a linha inicial da transação (método), tanto para a requisição do cliente quanto para a resposta do servidor, seguido de dois pontos (:) e um valor. Existem quatro tipos de cabeçalhos que poderão ser incluídos na mensagem os quais são: *general-header*, *requestheader*, *response-header* e *entity-header*.

Esses cabeçalhos são utilizados para enviar informações adicionais sobre a mensagem transmitida (*general-header*), a requisição e os clientes (*request-header*) que comunicam suas configurações e os formatos de documentos desejados como resposta. Além disso, são utilizados pelo servidor ao retornar o recurso no qual foi requisitado pelo cliente, para transmitir informações que descrevem as configurações do servidor e do recurso identificado pelo URI de requisição, e que não pertence à linha de status (*responseheader*). Na RFC 2616, estão descritos todos os campos que pertencem a esses cabeçalhos.

Uma mensagem HTTP pode conter um corpo de dados que são enviados abaixo das linhas de cabeçalho. Em uma mensagem de resposta, o corpo da mensagem é o recurso que foi requisitado pelo cliente, ou ainda uma mensagem de erro, caso este recurso não seja possível. Já em uma mensagem de requisição, o corpo pode conter dados que serão enviados diretamente pelo usuário ou um arquivo que será enviado para o servidor. Quando uma mensagem HTTP tiver um corpo, poderão ser incluídos cabeçalhos de entidades que descrevem suas características, como por exemplo, o *Content-Type* que informa o tipo MIME dos dados no corpo da mensagem e o *Content-Length* que informa a quantidade de bytes que o corpo da mensagem contém. A tabela ao lado apresenta alguns tipos MIME.

Exemplo	Descrição
text/plain	Arquivo no formato texto (ASCII)
text/html	Arquivo no formato HTML, utilizado como padrão para documentos Web
Image/gif	Imagem com o formato GIF
Image/jpeg	Imagem com o formato JPEG
application/zip	Arquivo compactado

De acordo com Fielding, uma mensagem de requisição do cliente é composta pelos seguintes campos: uma linha inicial (*Request-Line*); linhas de cabeçalhos (*Request-header*); uma linha em branco obrigatória e um corpo de mensagem opcional. A linha inicial de uma requisição é composta por três partes separadas por espaços: o método (*Method*), a identificação do URI (*Request-URI*) e a versão do HTTP (*HTTP-Version*) utilizado.

Segundo Bastos & Ladeira, *Request-URI* é um *identificador uniforme de recurso* (Uniform Resource Identifier) que identifica sobre qual recurso será aplicada a requisição. No protocolo HTTP, o tipo de URI utilizado é chamado de URL (Uniform Resource Locater), composto pela identificação do protocolo, pelo endereço do computador servidor e pelo documento requisitado.

O protocolo HTTP define oito métodos que indicam a ação a ser realizada no recurso especificado. Conforme Bastos e Ladeiras, o método determina o que o servidor deve fazer com o URL fornecido no momento da requisição de um recurso. Um servidor HTTP deve implementar ao menos os métodos GET e HEAD.

GET: Solicita algum recurso como um arquivo ou um *script* CGI (qualquer dado que estiver identificado pelo URI) por meio do protocolo HTTP. Por exemplo, segue abaixo uma comunicação entre um cliente e um servidor HTTP. O servidor possui a URL www.exemplo.com, porta 80.

O *pedido do cliente* (seguido por uma linha em branco, de maneira que o pedido termina com um *newline* duplo, cada um composto por um *carriage return* seguido de um *Line Feed*):

```
GET /index.html HTTP/1.1
Host: www.exemplo.com
```

O cabeçalho Host reconhece vários diferentes nomes DNS que tenham o mesmo IP.

A *resposta do servidor* (seguida por uma linha em branco e o texto da página solicitada):

```
HTTP/1.1 200 OK
Date: Mon, 23 May 2005 22:38:34 GMT
Server: Apache/1.3.27 (Unix)  (Red-Hat/Linux)
Last-Modified: Wed, 08 Jan 2003 23:11:55 GMT
Etag: "3f80f-1b6-3e1cb03b"
Content-Length: 438
Connection: close
Content-Type: text/html; charset=UTF-8
```

HEAD: Variação do GET em que o recurso não é retornado. É usado para obter metainformações por meio do cabeçalho da resposta, sem ter que recuperar todo o conteúdo.

POST: Envia dados para serem processados (por exemplo, dados de um formulário HTML) para o recurso especificado. Os dados são incluídos no corpo do comando. Sua utilização em uma requisição ocorre quando é necessário enviar dados ao servidor para serem processados, geralmente por um programa *script* identificado no *Request-URI*. Uma requisição por meio desse método sempre requer que as informações submetidas sejam incluídas no corpo da mensagem e formatadas como uma *query string*, além de conter cabeçalhos adicionais especificando seu tamanho (Content-Lenght) e seu formato (Content-Type). Por isso, esse método oferece uma maior segurança em relação aos dados transferidos, ao contrário do método GET que os dados são anexados a URL, ficando visíveis ao usuário. Por exemplo:

```
POST /index.html HTTP/1.0
Accept: text/html
If-modified-since: Sat, 29 Oct 1999 19:43:31 GMT
Content-Type: application/x-www-form-urlencoded
Content-Length: 30
Nome=NamePessoa&Idade=99&Curso=Computacao
```

PUT : Envia certo recurso.

DELETE: Exclui o recurso.

TRACE: Ecoa o pedido, de maneira que o cliente possa saber o que os servidores intermediários estão mudando em seu pedido.

OPTIONS: Recupera os métodos HTTP que o servidor aceita.

CONNECT: Serve para uso com um *proxy* que possa se tornar um túnel SSL (um túnel pode ser usado, por exemplo, para criar uma conexão segura).

Uma mensagem de resposta do servidor é composta pelos seguintes campos: uma linha inicial (*Status-Line*); linhas de cabeçalhos (*Responseheader*); uma linha em branco obrigatória e um corpo de mensagem opcional. A linha inicial de uma resposta, chamada de linha de status, possui por sua vez três partes separadas por espaços: a versão do protocolo HTTP (*HTTP-Version*), um código de status (*Status-Code*) da resposta, que fornece o resultado da requisição, e uma frase de justificativa (*Reason-Phrase*) que descreve o código do status.

A linha inicial de uma resposta HTTP indica ao cliente se sua requisição foi bem sucedida ou não. Essa situação é fornecida através de um código de retorno (*Status-Code*) e uma frase explicativa (*Reason-Phrase*). O código de status é formado por três dígitos e o primeiro dígito representa a classe que pertence classificada em cinco tipos:

- *1xx*: *Informational* (Informação) – utilizada para enviar informações para o cliente de que sua requisição foi recebida e está sendo processada;
- *2xx*: *Success* (Sucesso) – indica que a requisição do cliente foi bem sucedida;
- *3xx*: *Redirection* (Redirecionamento) – informa a ação adicional que deve ser tomada para completar a requisição;
- *4xx*: *Client Error* (Erro no cliente) – avisa que o cliente fez uma requisição que não pode ser atendida;
- *5xx*: *Server Error* (Erro no servidor) – ocorreu um erro no servidor ao cumprir uma requisição válida.

O protocolo HTTP define somente alguns códigos em cada classe descritos na RFC 2616, mas cada servidor pode definir seus próprios códigos.

O HTTP/1.0 é um protocolo sem estado. Isto significa que as conexões entre um cliente e um servidor são encerradas após o envio de cada requisição ou resposta. Cada vez que uma conexão é estabelecida ou encerrada, é consumida uma grande quantidade de tempo da CPU, de largura de banda e de memória.

Na maioria das vezes, para se obter o resultado esperado, é necessário realizar mais de uma solicitação de recursos através de várias conexões. Por exemplo, no caso de uma página Web, que consiste de diversos arquivos (.html, .gif, .css, etc) é preciso que sejam feitas várias requisições para compor a página, uma conexão não-persistente. O ideal seria que apenas uma conexão fosse utilizada para os pedidos e as respostas HTTP, diminuindo, assim, a sobrecarga ocasionada pelas conexões, uma conexão persistente.

A conexão persistente, implementada como conexão padrão no protocolo HTTP/1.1, possibilita que uma conexão seja estabelecida para enviar várias requisições em seqüência sem a necessidade de esperar por cada resposta, no qual serão recebidas na mesma ordem em que as solicitações foram enviadas, um processo chamado de *pipelining*. Pode também dar-se o caso de ser estabelecida uma conexão sem *pipelining*, em que o cliente só faz nova requisição quando o servidor lhe envia a resposta, ou seja, o servidor fica inactivo até o objecto (.html, .gif, .css, etc) atingir o seu destino no cliente.

Se uma requisição incluir o cabeçalho Connection: close, a conexão será encerrada após o envio da resposta correspondente. Utiliza-se este cabeçalho quando não há suporte a conexões persistentes, quando for a última requisição a ser enviada nesta conexão, ou ainda, sempre que quiser encerrar a conexão mesmo que nem todas as requisições tenham sido completadas. Além disso, o servidor pode fechar uma conexão se estiver ociosa por um determinado período de tempo.

9.7.8 SNMP

O **protocolo SNMP** (do inglês *Simple Network Management Protocol* - Protocolo Simples de Gerência de Rede) é um protocolo de gerência típica de redes TCP/IP, da camada de aplicação, que facilita o intercâmbio de informação entre os dispositivos de rede, como placas e comutadores (em inglês: *switches*). O SNMP possibilita aos administradores de rede gerenciar o desempenho da rede, encontrar e resolver seus eventuais problemas, e fornecer informações para o planejamento de sua expansão, dentre outras.

O software de gerência de redes **não** segue o modelo cliente-servidor convencional pois para as operações GET e SET a estação de gerenciamento se comporta como cliente e o dispositivo de rede a ser analisado ou monitorado se comporta como servidor, enquanto que na operação TRAP ocorre o oposto, pois no envio de alarmes é o dispositivo gerenciado que toma iniciativa da comunicação. Por conta disso, os sistemas de gerência de redes evitam os termos 'cliente' e 'servidor' e optam por usar "gerente" para a aplicação que roda na estação de gerenciamento e "agente" para a aplicação que roda no dispositivo de rede.

O programa gerente da rede é a entidade responsável pelo monitoramento e controle dos sistemas de hardware e software que compõem a rede, e o seu trabalho consiste em detectar e corrigir problemas que causem ineficiência (ou impossibilidade) na comunicação e eliminar as condições que poderão levar a que o problema volte a surgir.

A gerência de uma rede pode não ser simples, dada sua heterogeneidade em termos de hardware e software, e de componentes da rede, por vezes incompatíveis. As falhas intermitentes, se não forem detectadas, podem afetar o desempenho da rede. Um software de gerência de redes permite ao gestor monitorar e controlar os componentes da sua rede.

Uma rede gerenciada pelo protocolo SNMP é formada por três componentes chaves:

1. Dispositivos Gerenciados
2. Agentes
3. Sistemas de Gerenciamento de Redes (NMS - *Network-Management Systems*)

Um Dispositivo Gerenciado é um nó de rede que possui um agente SNMP instalado e se encontra em uma rede gerenciada. Estes dispositivos coletam e armazenam informações de gerenciamento e mantém estas informações disponíveis para sistemas NMS através do protocolo SNMP. Dispositivos gerenciados, também às vezes denominados de dispositivos de rede, podem ser roteadores, servidores de acesso, impressoras, computadores, servidores de rede, switches, dispositivos de armazenamento, dentre outros.

Um Agente é um módulo de software de gerenciamento de rede que fica armazenado em um Dispositivo Gerenciado. Um agente tem o conhecimento das informações de gerenciamento locais e traduz estas informações para um formato compatível com o protocolo SNMP.

Um sistema NMS é responsável pelas aplicações que monitoram e controlam os Dispositivos Gerenciados. Normalmente é instalado em um (ou mais de um) servidor de rede dedicado a estas operações de gerenciamento, que recebe informações (pacotes SNMP) de todos os dispositivos gerenciados daquela rede.

O framework SNMP consiste de: Agentes Mestres (Master Agents), Sub-agentes (Subagents) e Estações de Gerenciamento (Management Stations).

O Master Agent em uma rede gerenciada é, na verdade, um software sendo executado em um dispositivo com suporte a SNMP, por exemplo, um roteador, que interage com uma estação de gerenciamento. É o equivalente a um servidor, na comunicação cliente/servidor, ou a um daemon, sob o ponto de vista de sistemas operacionais. Os subagentes são os responsáveis por passarem informações específicas para o Masters Agent.

Os subagentes ou *subagents* são pequenos programas em execução no dispositivo com suporte a SNMP, responsáveis pelo monitoramento de recursos específicos naquele dispositivo, como por exemplo, o status de um link ethernet em um roteador, ou a quantidade de espaço livre em um disco de um servidor. Algumas características dos softwares subagentes são:

- Coletar informações de objetos gerenciados
- Configurar parâmetros destes objetos gerenciados
- Responder a solicitações do software de gerência da rede
- Gerar alarmes ou *traps* em determinadas situações

O Gerente da Rede ou Estação de Gerenciamento ou ainda *Management Station* é o componente final da arquitetura de uma solução SNMP. Funciona como um cliente em uma comunicação cliente/servidor. Realiza requisições de informações aos dispositivos gerenciados, que podem ser temporárias ou através de comandos a qualquer tempo. E ainda é o responsável por receber alarmes gerados pelos agentes e gerar saídas para estes alarmes, tais como, alterar (SET) o valor de um determinado parâmetro gerenciado no equipamento, enviar mensagem para o celular do administrador da rede, dentre outras.

O SNMP é um protocolo padrão usado para gerência de redes, que define os formatos dos pedidos que o *Gerente* envia para o *Agente* e os formatos das respostas que o *agente* retorna, assim como o significado exato de cada pedido e resposta. Uma mensagem SNMP é codificada com um padrão designado de ASN.1 (do inglês: *Abstract Syntax Notation.1*).

O ASN.1 para permitir a transferência de grandes pacotes, sem desperdiçar espaço em cada transferência, usa uma combinação de tamanho e valor para cada objeto a ser transferido.

O SNMP não define um grande número de comandos, em lugar disso define duas operações básicas:

- *GET*, para obter um valor de um dispositivo
- *SET*, para colocar um valor num dispositivo

O comando que especifica uma operação de *GET* ou *SET* deve especificar o nome do objeto, que é único.

Podemos definir objetos. No caso de um contador de erros de CRC e uma vez que o SNMP não inclui comandos específicos para fazer *reset* do contador, uma forma simples é colocar zero no contador. Neste caso, o Gerente faz o *GET* (leitura) do parâmetro desejado para determinar o estado do dispositivo. As operações que controlam o dispositivo são definidas como efeitos secundários de *SET* (alterar/gravar valores) em objetos.

[[Especifica (na versão 1) quatro unidades de dados do protocolo (PDU):

1. GET, usado para retirar um pedaço de informação de gerenciamento.
2. GETNEXT, usado interativamente para retirar sequências de informação de gerenciamento.
3. SET, usado para fazer uma mudança no subsistema gerido.
4. TRAP, usado para reportar uma notificação ou para outros eventos assíncronos sobre o subsistema gerido.]]

Todos os objetos acessados pelo SNMP devem ter nomes únicos definidos e atribuídos. Além disso, o *Gerente* e o *Agente* devem acordar os nomes e significados das operações *GET* e *SET*. O conjunto de todos os objetos SNMP é coletivamente conhecido como MIB (do inglês: Management Information Base). O standard SNMP não define o MIB, mas apenas o formato e o tipo de codificação das mensagens. A especificação das variáveis MIB, assim como o significado das operações *GET* e *SET* em cada variável, são especificados por um padrão próprio.

A definição dos objetos do MIB é feita com o esquema de nomes do ASN.1, o qual atribui a cada objeto um prefixo longo que garante a unicidade do nome, a cada nome é atribuído um número inteiro. Também, o SNMP não especifica um conjunto de variáveis, e como a definição de objetos é independente do protocolo de comunicação, permite criar novos conjuntos de variáveis MIB, definidos como standards, para novos dispositivos ou novos protocolos. Por isso, foram criados muitos conjuntos de variáveis MIB que correspondem a protocolos como UDP, IP, ARP, assim como variáveis MIB para hardware de rede como Ethernet ou FDDI, ou para dispositivos tais como bridges, switches ou impressoras.

A versão 2 do SNMP é uma evolução do protocolo inicial. O SNMPv2 oferece uma boa quantidade de melhoramentos em relação ao SNMPv1, incluindo operações adicionais do protocolo, melhoria na performance, segurança, confidencialidade e comunicações Gerente-para-Gerente. A padronização de uma outra versão do SNMP - o SNMPv3 ainda está em desenvolvimento, definido nos RFC 3411 -RFC 3418.

Na prática, as implementações do SNMP oferecem suporte para as múltiplas versões (RFC 3584), tipicamente SNMPv1, SNMPv2c e SNMPv3.

9.7.9 BOOTP e DHCP

No inicio o BOOTP permitia a configuração automática de impressoras e hosts de uma rede. Ele fazia isso associando um numero MAC a um endereço IP - ou outro tipo de configuração. Com a evolução das redes o protocolo BOOTP ficou defasado e estava encontrando problemas para a configuração automática em grandes redes. Isso desencadeou uma série de estudos e um grupo denominado IETF - Internet

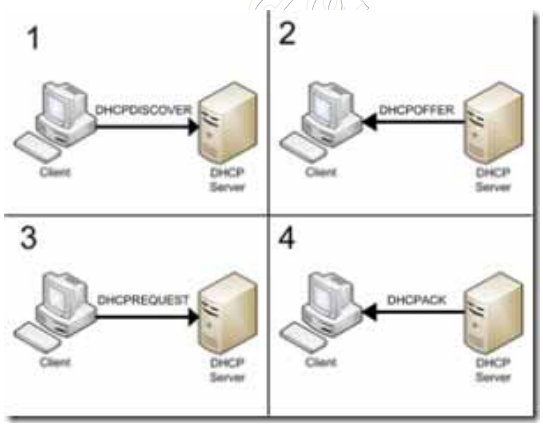
Engineering Task Force - que desenvolveu um protocolo que substituiria o BOOTP, sanando suas limitações, o DHCP.

O DHCP possibilita a configuração automática de hosts numa rede. Isso facilita muito o trabalho de um profissional que esteja fazendo a configuração da rede. Por exemplo, numa rede de poucos computadores a configuração máquina a máquina na rede não é um trabalho tão árduo. Porém, se a rede for muito extensa - 100 ou mais computadores - esse trabalho se torna um verdadeiro calvário. Com o DHCP atuando, esse trabalho é reduzido. Ele faz a distribuição de IPs de acordo com as requisições dos hosts que entram na rede.

Outro bom exemplo do uso do DHCP são os provedores de acesso a internet. Em sua maioria o cliente - usuário - recebe um ip diferente para cada nova conexão. Isso torna-se possível com a aplicação híbrida do DHCP com um serviço de PPP - Point to Point Protocol, por exemplo. Sem o DHCP um ou mais funcionários da empresa provedora de acesso deveriam visitar os clientes para configurar a conexão. E a cada problema que o serviço viesse a ter esses mesmos funcionários deveriam visitar novamente os clientes. A provedora de acesso teria mais custos configurando as maquinas dos clientes do que lucro com o serviço prestado.

O **DHCP**, *Dynamic Host Configuration Protocol*, é um protocolo de serviço TCP/IP que oferece configuração dinâmica de terminais, com concessão de endereços IP de host e outros parâmetros de configuração para clientes de rede. Este protocolo é o sucessor do BOOTP que, embora mais simples, tornou-se limitado para as exigências atuais. O DHCP surgiu como standard em Outubro de 1993. O RFC 2131 contém as especificações mais atuais (Março de 1997). O último standard para a especificação do DHCP sobre IPv6 (DHCPv6) foi publicado a Julho de 2003 como RFC 3315.

Resumidamente, o DHCP opera da seguinte forma:



- Um cliente envia um pacote UDP em *broadcast* (destinado a todas as máquinas) com um pedido DHCP;
- Os servidores DHCP que capturarem este pacote irão responder (se o cliente se enquadrar numa série de critérios — ver abaixo) com um pacote com configurações onde constará, pelo menos, um endereço IP, uma máscara de rede e outros dados opcionais, como o gateway, servidores de DNS, etc.

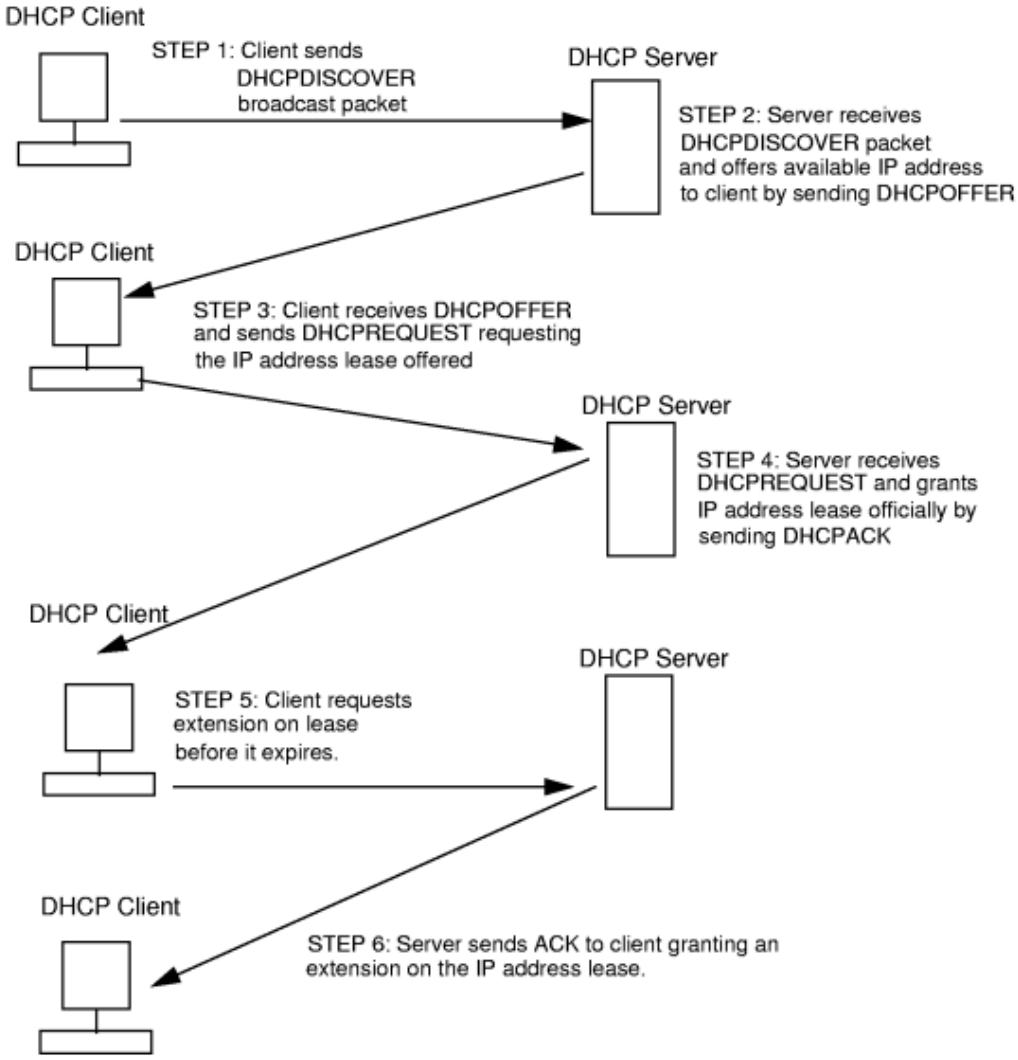
O DHCP usa um modelo cliente-servidor, no qual o servidor DHCP mantém o gerenciamento centralizado dos endereços IP usados na rede.

Detalhadamente, o DHCP opera da seguinte forma:

Para encontrar seu endereço IP, uma máquina recém-inicializada numa rede com um servidor DHCP realiza por padrão os seguintes passos:

1. A maquina cliente transmite via broadcast um pacote UDP chamado 'DHCP DISCOVER'.
2. O servidor recebe a requisição e oferece um IP através de um pacote 'DHCP OFFER'.
3. O novo cliente recebe a oferta do IP e solicita um 'lease'[1] do IP ofertado através do pacote 'DHCP REQUEST'.

4. O servidor recebe a confirmação e garante oficialmente o 'lease' do IP ao cliente através do pacote 'DHCP ACK'.
5. Se o cliente ainda estiver utilizando o IP e o tempo de 'lease' estiver acabando ele solicita ao servidor que o tempo do 'lease' seja estendido.
6. O servidor envia um 'ACK' garantindo o prolongamento do prazo de utilização do IP pelo cliente.



O DHCP oferece três tipos de alocação de endereços IP:

• **Atribuição manual** - Onde existe uma tabela de associação entre o Endereço MAC do cliente (que será comparado através do pacote *broadcast* recebido) e o endereço IP (e restantes dados) a fornecer. Esta associação é feita manualmente pelo administrador de rede; por conseguinte, apenas os clientes cujo MAC consta nesta lista poderão receber configurações desse servidor;

• **Atribuição automática** - Onde o cliente obtém um endereço de um espaço de endereços possíveis, especificado pelo administrador. Geralmente não existe vínculo entre os vários MAC habilitados a esse espaço de endereços;

- **Atribuição dinâmica** - O único método que dispõe a reutilização dinâmica dos endereços. O administrador disponibiliza um espaço de endereços possíveis, e cada cliente terá o software TCP/IP da sua interface de rede configurados para requisitar um endereço por DHCP assim que a máquina arranque. A alocação utiliza um mecanismo de *aluguer* do endereço, caracterizado por um tempo de vida. Após a máquina se desligar, o tempo de vida naturalmente irá expirar, e da próxima vez que o cliente se ligue, o endereço provavelmente será outro.

Algumas implementações do software servidor de DHCP permitem ainda a actualização dinâmica dos servidores de DNS para que cada cliente disponha também de um DNS. Este mecanismo utiliza o protocolo de atualização do DNS especificado no RFC 2136

9.7.10 TLS e SSL

O **Transport Layer Security** - TLS (em português: *Protocolo de Camada de Sockets Segura*) e o seu predecessor, **Secure Sockets Layer** - SSL, são protocolos criptográficos que provêem comunicação segura na Internet para serviços como email (SMTP), navegação por páginas (HTTP) e outros tipos de transferência de dados. Há algumas pequenas diferenças entre o SSL 3.0 e o TLS 1.0, mas o protocolo permanece substancialmente o mesmo. O termo "SSL" usado aqui aplica-se a ambos os protocolos, exceto se disposto em contrário. O protocolo SSL 3.0 também é conhecido como SSL3, e o TLS 1.0 como TLS1 ou ainda SSL3.1.

Baseia-se no protocolo TCP da suíte TCP/IP e utiliza-se do conceito introduzido por Diffie-Hellman nos anos 70 (criptografia de chave pública) e Phil Zimmerman (criador do conceito PGP).

Desenvolvido pela Netscape, o SSL versão 3.0 foi lançado em 1996, e serviu posteriormente de base para o desenvolvimento do TLS versão 1.0, um protocolo padronizado da IETF originalmente definido pelo RFC 2246. Grandes instituições financeiras como Visa, MasterCard, American Express, dentre outras, aprovaram o SSL para comércio eletrônico seguro na Internet.

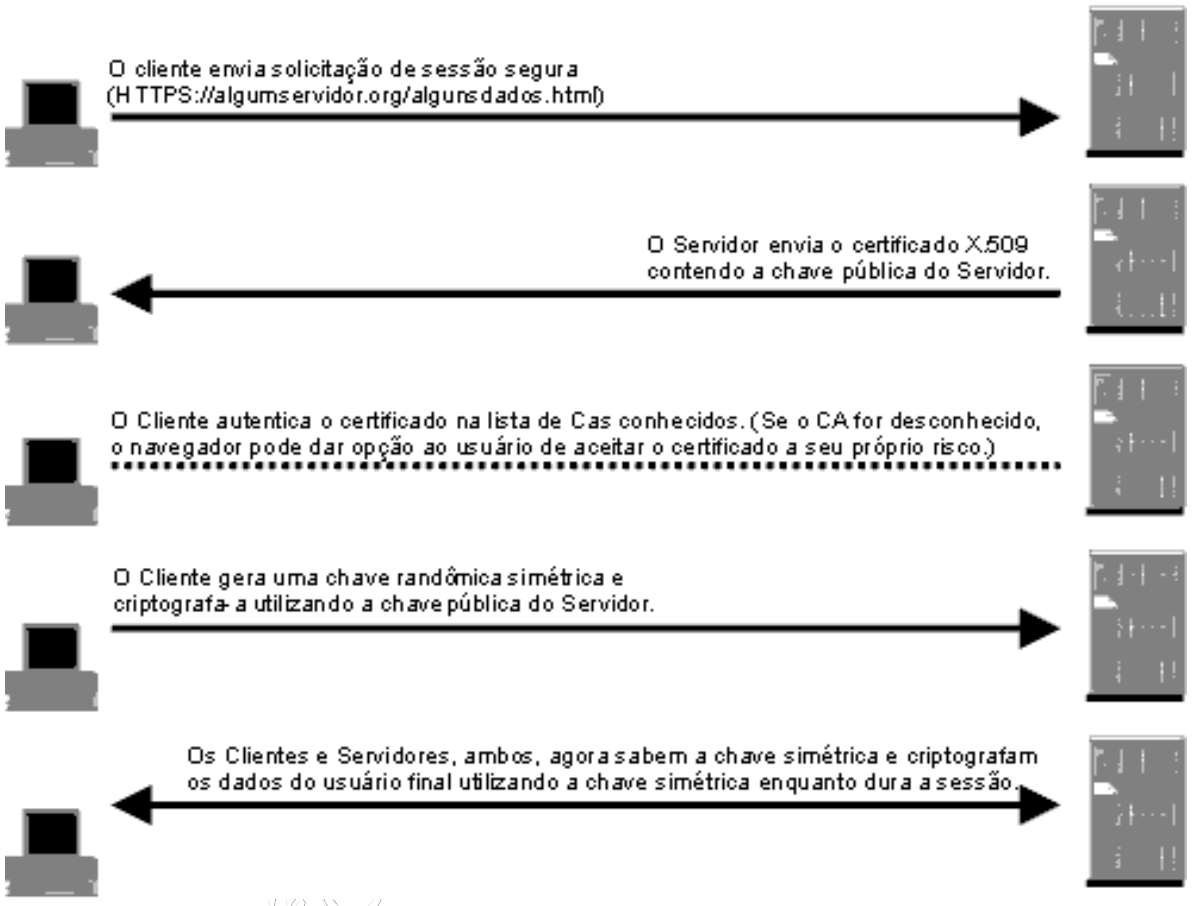
O SSL opera de forma modular, possui *design* extensível, e apresenta compatibilidade entre pares com versões diferentes do mesmo.

O SSL é um protocolo que fornece privacidade e integridade entre os dois aplicativos de comunicação, utilizando TCP/IP. O *Hypertext Transfer Protocol* (HTTP) para a World Wide Web utiliza SSL para executar comunicações seguras.

Os dados que vem e voltam entre o cliente e o servidor são criptografados utilizando um algoritmo simétrico, como DES ou RC4. Um algoritmo de chave pública -normalmente RSA- é utilizado para trocar as chaves criptografadas e para as assinaturas digitais. O algoritmo utiliza a chave pública no certificado digital do servidor. Com o certificado digital do servidor, o cliente pode verificar a identidade do servidor. As versões 1 e 2 do protocolo SSL fornecem somente autenticação de servidor. A versão 3 inclui autenticação de cliente, utilizando os certificados digitais do servidor e do cliente.

Uma conexão SSL é iniciada sempre pelo cliente. No início da sessão do SSL, um protocolo de reconhecimento SSL é executado. Esse protocolo de reconhecimento produz os parâmetros criptográficos da sessão. Uma visão geral simplificada de como o protocolo de reconhecimento é processado será

mostrada abaixo. Esse exemplo assume que a conexão SSL está sendo estabelecida entre um navegador da Web e um servidor da Web.



1. O cliente envia uma mensagem "olá" ao cliente que lista as capacidades criptográficas do cliente (classificada em ordem de preferência do cliente), como a versão do SSL, os cipher suites suportados pelo cliente e os métodos de compactação de dados suportados pelo cliente. A mensagem também contém um número aleatório de 28 bytes.
2. O servidor responde a mensagem "olá" do servidor que contém o método criptográfico (cipher suite) e o método de compactação de dados selecionado pelo servidor, o ID da sessão e outros números aleatórios.

   Nota: *O cliente e o servidor deve suportar pelo menos um cipher suite (conjunto de cifras criptográficas) comum, caso contrário o protocolo de reconhecimento falhará. O servidor escolhe normalmente o cipher suite comum mais resistente.*
3. O servidor envia seu certificado digital. (O servidor utiliza certificados digitais X.509 V3 com SSL.). Se o servidor utilizar SSL V3 e ser a aplicação servidor (por exemplo, o servidor da Web) exigir um certificado digital para autenticação do cliente, o servidor enviará uma mensagem de "pedido de certificado digital". Na mensagem "pedido de certificado digital", o servidor envia uma lista de tipos de certificados digitais suportados e os nomes distintos de autoridades de certificação aceitáveis.
4. O servidor envia uma mensagem de servidor "olá enviado" e aguarda uma resposta do cliente.
5. Ao receber a mensagem de servidor "olá enviado", o cliente (navegador da Web) verifica a validade do certificado digital do servidor e verifica se os parâmetros de "olá" no servidor são aceitos. Se o servidor

solicitar um certificado digital do cliente, o cliente enviará um certificado digital e, se nenhum certificado digital apropriado estiver disponível, o cliente enviará um alerta "nenhum certificado digital". Esse alerta é apenas um aviso, mas a aplicação do servidor pode causar falha na sessão se a autenticação do cliente for obrigatória.

6. O cliente envia uma mensagem "intercâmbio de chave do cliente". Essa mensagem contém o *segredo pré-master*, um número aleatório de 46 bytes utilizado na geração de chaves de criptografia simétrica e chaves de MAC (*códigos de autenticação de mensagens*), criptografados com a chave pública do servidor. Se o cliente enviar um certificado digital para o servidor, o cliente envia uma mensagem "verificação de certificado digital" assinado com a chave privativa do cliente. Verificando a assinatura dessa mensagem, o servidor pode analisar explicitamente a propriedade do certificado digital do cliente. ***Nota:*** *Um processo adicional para verificar o certificado digital não é necessário. Se o servidor não tiver uma chave privativa que pertença ao certificado digital, ele não poderá descriptografar o segredo pré-master e criar as chaves corretas para o algoritmo de criptografia simétrica, e o protocolo de reconhecimento irá falhar.*
7. O cliente utiliza uma série de operações criptográficas para converter um segredo pré-master em um *segredo master*, a partir do qual todos os materiais de chave exigidos para criptografia e para autenticação de mensagem é derivado. Então o cliente envia uma mensagem "alterar cipher spec" para para fazer o servidor ir para o cipher suite negociado mais recente. A mensagem de texto enviada pelo cliente (a mensagem "concluído") é a primeira mensagem criptografada com esse método cipher e essas chaves.
8. O servidor responde a mensagem "alterar cipher spec" e "concluído" por si só.
9. O protocolo de reconhecimento SSL é concluído e os dados do aplicativo criptografados podem ser enviados.

O Secure Sockets Layer V3 pode utilizar os certificados digitais do servidor bem como as do cliente. Como explicado anteriormente, os certificados digitais do servidor são obrigatórias para uma sessão SSL, enquanto os certificados digitais do cliente são opcionais, dependendo dos requisitos de autenticação do cliente.

O *PKI* (public key infrastructure) utilizado pelo SSL permite quaisquer números de autoridades de certificação de raiz. Uma organização ou um usuário final deve decidir quais CAs serão aceitas como *confiáveis*. Para poder verificar os certificados digitais do servidor, o cliente deve possuir os certificados digitais de raiz CA utilizados pelo servidor.

Se uma sessão SSL está para ser estabelecida com um servidor que envia um certificado digital com raiz CA que não está definido no arquivo truststore de cliente, a sessão SSL não será estabelecida. Para evitar essa situação, importe o certificado digital de raiz CA para o armazenamento de chave ou o truststore de cliente.

Se a autenticação do cliente for utilizada, o servidor exigirá a posse dos certificados digitais de raiz CA utilizados pelos clientes. Todos os certificados digitais de raiz CA que não fazem parte do armazenamento de chave de servidor padrão devem ser instalados usando o utilitário iKeyman antes da emissão dos certificados digitais do cliente feita por CAs.

10 COMPETÊNCIA 5 – PRÁTICA DE CABEAMENTO EM REDES

10.1 CRIMPAGEM DE CABOS DIRETOS E INVERTIDOS

A ferramenta básica para crimpar os cabos é o alicate de crimpagem. Ele "esmaga" os contatos do conector, fazendo com que as facas-contato perfurem a cobertura plástica e façam contato com os fios do cabo de rede:

É possível comprar alicates de crimpagem razoáveis por pouco mais de 50 reais, mas existem alicates de crimpagem para uso profissional que custam bem mais. Existem ainda "alicates" mais baratos, com o corpo feito de plástico, que são mais baratos, mas não valem o papelão da embalagem. Alicates de crimpagem precisam ser fortes e precisos, por isso evite produtos muito baratos.

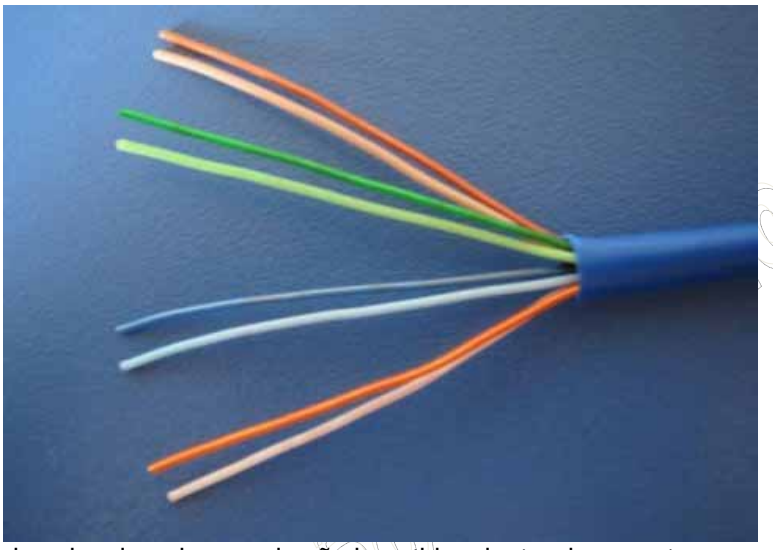
Antes de iniciar, observe o alicate de crimpagem. Nele existem dois tipos de guilhotinas: uma para desencapar os cabos e outra para aparar os fios. Em alguns casos existe um sulco no qual o cabo deve ser inserido para ser descascado. Existe também um conector no qual serão crimpados os conectores RJ-45.

Ao crimpar os cabos de rede, o primeiro passo é descascar os cabos, tomando cuidado para não ferir os fios internos, que são bastante finos. Normalmente, o alicate inclui uma saliência no canto da guilhotina, que serve bem para isso. Existem também descascadores de cabos específicos para cabos de rede, que são sempre um item bem-vindo na caixa de ferramentas:

Os quatro pares do cabo são diferenciados por cores. Um par é laranja, outro é azul, outro é verde e o último é marrom. Um dos cabos de cada par tem uma cor sólida e o outro é mais claro ou malhado, misturando a cor e pontos de branco. É pelas cores que diferenciamos os 8 fios.

O segundo passo é destrançar os cabos, deixando-os soltos. Para facilitar o trabalho, descasque um pedaço grande do cabo, uns 5 ou 6 centímetros, para poder organizar os cabos com mais facilidade e depois corte o excesso, deixando apenas a meia polegada de cabo (1.27 cm, ou menos) que entrará dentro do conector.

O próprio alicate de crimpagem inclui uma guilhotina para cortar os cabos, mas operá-la exige um pouco de prática, pois você precisa segurar o cabo com uma das mãos, mantendo os fios na ordem correta e manejar o alicate com a outra. A guilhotina faz um corte reto, deixando os fios prontos para serem inseridos dentro do conector, você só precisa mantê-los firmes enquanto encaixa e crimpa o conector.



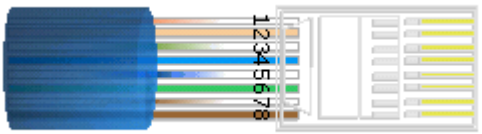
Existem dois padrões para a ordem dos fios dentro do conector, o EIA 568B (o mais comum) e o EIA 568A. A diferença entre os dois é que a posição dos pares de cabos laranja e verde são invertidos dentro do conector.

Existe muita discussão em relação com qual dos dois é "melhor", mas na prática não existe diferença de conectividade entre os dois padrões. A única observação é que você deve cabear toda a rede utilizando o mesmo padrão. Como o EIA 568B é de longe o mais comum, recomendo que você o utilize ao crimpar seus próprios cabos.

Uma observação é que muitos cabos são certificados para apenas um dos dois padrões; caso encontre instruções referentes a isso nas especificações, ou decalcadas no próprio cabo, crimpe os cabos usando o padrão indicado.

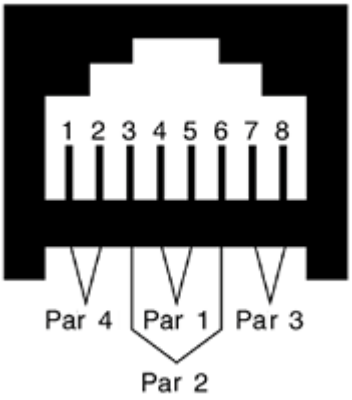
No padrão EIA 568B, a ordem dos fios dentro do conector (em ambos os lados do cabo) é a seguinte:

1- Branco com Laranja
2- Laranja
3- Branco com Verde
4- Azul
5- Branco com Azul
6- Verde
7- Branco com Marrom
8- Marrom



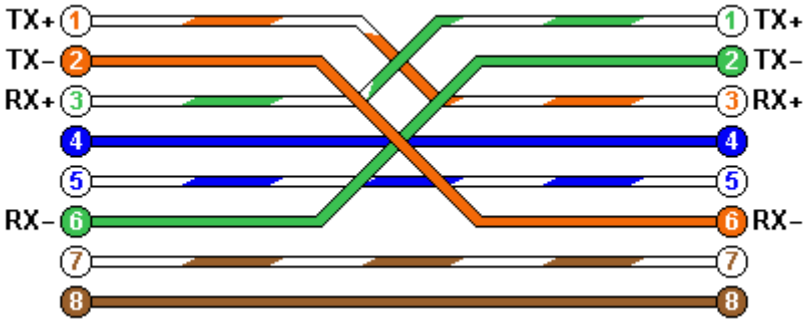
Os cabos são encaixados nessa ordem, com a trava do conector virada para baixo, como no diagrama.

Ou seja, se você olhar o conector "de cima", vendo a trava, o par de fios laranja estará à direita e, se olhar o conector "de baixo", vendo os contatos, eles estarão à esquerda. Este outro diagrama mostra melhor como fica a posição dos cabos dentro do conector:



O cabo crimpado com a mesma disposição de fios em ambos os lados do cabo é chamado de cabo "reto", ou **straight**. Este é o tipo "normal" de cabo, usado para ligar os micros ao switch ou ao roteador da rede. Existe ainda um outro tipo de cabo, chamado de "**cross-over**" (também chamado de cabo cross, ou cabo cruzado), que permite ligar diretamente dois micros, sem precisar do hub ou switch. Ele é uma opção mais barata quando você tem apenas dois micros.

No cabo cruzado, a posição dos fios é diferente nos dois conectores, de forma que o par usado para enviar dados (TX) seja ligado na posição de recepção (RX) do segundo micro e vice-versa. De um dos lados a pinagem é a mesma de um cabo de rede normal, enquanto no outro a posição dos pares verde e laranja são trocados. Daí vem o nome cross-over, que significa, literalmente, "cruzado na ponta":



Para fazer um cabo cross-over, você crimpa uma das pontas seguindo o padrão EIA 568B que vimos acima e a outra utilizando o padrão EIA 568A, onde são trocadas as posições dos pares verde e laranja:

1- Branco com Verde
2- Verde
3- Branco com Laranja
4- Azul
5- Branco com Azul
6- Laranja
7- Branco com Marrom
8- Marrom

A maioria dos switches atuais são capazes de "descruzar" (*autosense*) os cabos automaticamente quando necessário, permitindo que você misture cabos normais e cabos cross-over dentro do cabeamento da rede. Graças a isso, a rede vai funcionar mesmo que você use um cabo cross-over para conectar um dos micros ao hub por engano.

Este cabo cross-over "clássico" pode ser usado para ligar placas de 10 ou 100 megabits, onde as transmissões são na realidade feitas usando apenas dois dos pares dos cabos. Placas e switches Gigabit Ethernet utilizam os quatro pares e por isso precisam de um cabo cross-over especial, crimpado com uma pinagem diferente. Usando um cabo cross convencional, a rede até funciona, mas as placas são forçadas a reduzir a velocidade de transmissão para 100 megabits, de forma a se adaptarem ao cabeamento.

Para fazer um cabo cross-over Gigabit Ethernet, você deve utilizar o padrão EIA 568B (Branco com Laranja, Laranja, Branco com Verde, Azul, Branco com Azul, Verde, Branco com Marrom, Marrom) de um dos lados do cabo, como usaria ao crimpar um cabo normal. A mudança vem ao crimpar o outro lado do cabo, onde é usada a seguinte pinagem:

1- Branco com Verde
2- Verde
3- Branco com Laranja
4- Branco com Marrom
5- Marrom
6- Laranja
7- Azul
8- Branco com Azul

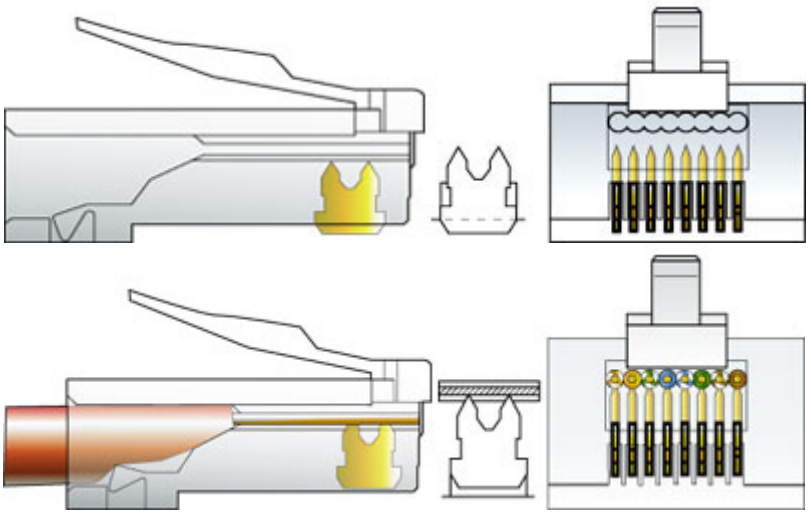
Muitos switches e também algumas placas Gigabit podem ser ligados diretamente usando cabos *straight*, pois os transmissores são capazes de ajustar a transmissão via software, recurso chamado de Auto-MDI/MDI-X. Entretanto, nem todos os dispositivos suportam o recurso, de forma que os cabos cross-over ainda são necessários em diversas situações. Revisando, os padrões para os três tipos de cabos são:

Cabo straight (10, 100 ou 1000 megabits):	
1- Branco com Laranja	1- Branco com Laranja
2- Laranja	2- Laranja
3- Branco com Verde	3- Branco com Verde
4- Azul	4- Azul
5- Branco com Azul	5- Branco com Azul
6- Verde	6- Verde
7- Branco com Marrom	7- Branco com Marrom
8- Marrom	8- Marrom
Cabo cross-over (10 ou 100 megabits):	
1- Branco com Laranja	1- Branco com Verde
2- Laranja	2- Verde
3- Branco com Verde	3- Branco com Laranja
4- Azul	4- Azul
5- Branco com Azul	5- Branco com Azul
6- Verde	6- Laranja
7- Branco com Marrom	7- Branco com Marrom
8- Marrom	8- Marrom

Cabo cross-over para Gigabit Ethernet	
1- Branco com Laranja	1- Branco com Verde
2- Laranja	2- Verde
3- Branco com Verde	3- Branco com Laranja
4- Azul	4- Branco com Marrom
5- Branco com Azul	5- Marrom
6- Verde	6- Laranja
7- Branco com Marrom	7- Azul
8- Marrom	8- Branco com Azul

Ao crimpar, você deve retirar apenas a capa externa do cabo e não descascar individualmente os fios, pois isso, ao invés de ajudar, serviria apenas para causar mau contato, deixando frouxo o encaixe com os pinos do conector.

A função do alicate é fornecer pressão suficiente para que os pinos do conector RJ-45, que internamente possuem a forma de lâminas, esmaguem os fios do cabo, alcançando o fio de cobre e criando o contato:

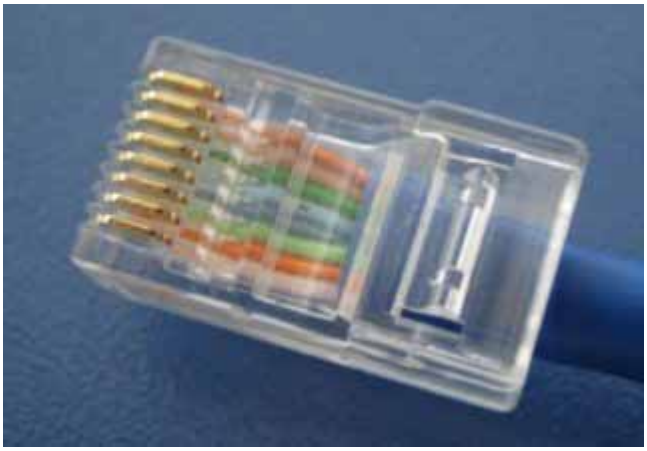


Como os fios dos cabos de rede são bastante duros, é preciso uma boa dose de força para que o conector fique firme, daí a necessidade de usar um alicate resistente. Não tenha medo de quebrar ou danificar o alicate ao crimpar, use toda a sua força:

É preciso um pouco de atenção ao cortar e encaixar os fios dentro do conector, pois eles precisam ficar perfeitamente retos. Isso demanda um pouco de prática. No começo, você vai sempre errar algumas vezes antes de conseguir.

Veja que o que protege os cabos contra as interferências externas são justamente as tranças. A parte destrançada que entra no conector é o ponto fraco do cabo, onde ele é mais

vulnerável a todo tipo de interferência. Por isso, é recomendável deixar o menor espaço possível sem as tranças. Para crimpar cabos dentro do padrão, você precisa deixar menos de meia polegada de cabo (1.27 cm) destrançado. Você só vai conseguir isso cortando o excesso de cabo solto antes de encaixar o conector, como na foto:



Outra observação é que, além de ser preso pelos conectores metálicos, o cabo é preso dentro do conector através de uma trava plástica, que é também presa ao crimpar o cabo. A trava prende o cabo através da cobertura plástica, por isso é importante cortar todo o excesso de cabo destrançado, fazendo com que parte da cobertura plástica fique dentro do conector e seja presa pela trava. Sem isso, os contatos podem facilmente ser rompidos com qualquer esbarrão, tornando a rede como um todo menos confiável.

Além do cabo e do conector RJ-45, existem dois acessórios, que você pode ou não usar em seus cabos, conforme a disponibilidade. O primeiro são as capas plásticas (boots), que são usadas nas pontas dos cabos para melhorar o aspecto visual. Por estarem disponíveis em várias cores, elas podem ser também usadas para identificar os cabos, mas com exceção disso elas são puramente decorativas, não possuem nenhuma outra função. Para usá-las, basta colocar a capa antes do conector:

O segundo são os inserts, que são um tipo de suporte plástico que vai dentro do conector. Depois de destrançar, organizar e cortar o excesso de cabo, você passa os 8 fios dentro do insert e eles os mantêm na posição, facilitando o encaixe no conector.

Os conectores RJ-45 projetados para uso em conjunto com o insert possuem um espaço interno maior para acomodá-lo. Devido a isso, os inserts são fornecidos em conjunto com alguns modelos de conectores e raramente são vendidos separadamente:

O primeiro teste para ver se os cabos foram crimpados corretamente é conectar um dos micros (ligado) ao switch e ver se os LEDs da placas de rede e do hub acendem. Isso mostra que os sinais elétricos enviados estão chegando até o switch e que ele foi capaz de abrir um canal de comunicação com a placa.

Se os LEDs nem acenderem, então não existe o que fazer. Corte os conectores e tente de novo. Infelizmente, os conectores são descartáveis: depois de crimpar errado uma vez, você precisa usar outro novo, aproveitando apenas o cabo. Mais um motivo para prestar atenção ;).

Existem também aparelhos testadores de cabos, que oferecem um diagnóstico muito mais sofisticado, dizendo, por exemplo, se os cabos são adequados para transmissões a 100 ou a 1000 megabits e avisando caso algum dos 8 fios do cabo esteja rompido. Os mais sofisticados avisam inclusive em que ponto o cabo está rompido, permitindo que você aproveite a parte boa.

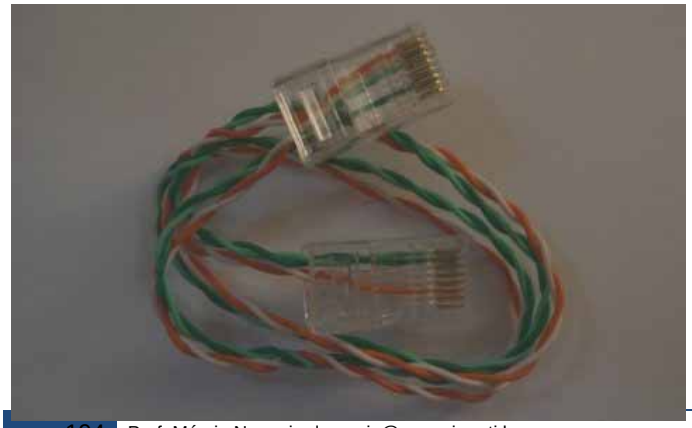
10.2 TESTE DE CABOS

Esses aparelhos serão bastante úteis se você for crimpar muitos cabos, mas são dispensáveis para trabalhos esporádicos, pois é muito raro que os cabos venham com fios rompidos de fábrica. Os cabos de rede apresentam também uma boa resistência mecânica e flexibilidade, para que possam passar por dentro de tubulações. Quase sempre os problemas de transmissão surgem por causa de conectores mal crimpados.

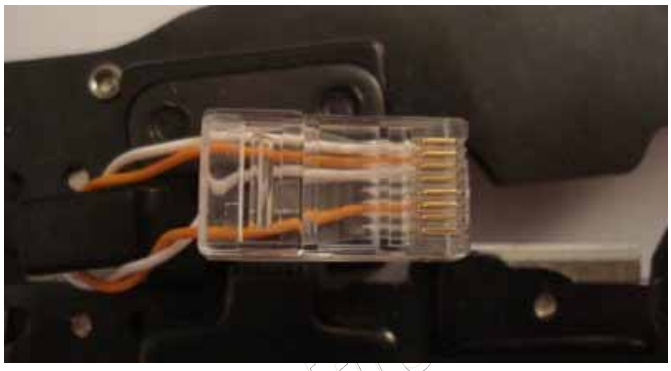
Existem ainda modelos mais simples de testadores de cabos, que chegam a custar em torno de 20 reais. Eles realizam apenas um teste de continuidade do cabo, checando se o sinal elétrico chega até a outra ponta e, verificando o nível de atenuação, para certificar-se de que ele cumpre as especificações mínimas. Um conjunto de 8 leds se acende, mostrando o status de cada um dos 8 fios. Se algum fica apagado durante o teste, você sabe que o fio correspondente está partido. A limitação é que eles não são capazes de calcular em que ponto o cabo está partido, de forma que a sua única opção acaba sendo trocar e descartar o cabo inteiro.

Uma curiosidade com relação aos testadores é que algumas placas-mãe da Asus, com rede Yukon Marvel (e, eventualmente, outros modelos lançados futuramente), incluem um software testador de cabos, que pode ser acessado pelo setup, ou através de uma interface dentro do Windows. Ele funciona de uma forma bastante engenhosa. Quando o cabo está partido em algum ponto, o sinal elétrico percorre o cabo até o ponto onde ele está rompido e, por não ter para onde ir, retorna na forma de interferência. O software cronometra o tempo que o sinal demora para ir e voltar, apontando com uma certa precisão depois de quantos metros o cabo está rompido.

Outra dica é que no padrão 100BASE-TX são usados apenas os pares laranja e verde para transmitir dados. Você pode tirar proveito disso para fazer um cabo mini-crossover para levar na sua caixa de ferramentas, usando apenas os pares laranja e verde do cabo. De um lado a pinagem seria: branco com laranja, laranja, branco com verde, nada, nada, verde, nada, nada; e do outro seria: branco com verde, verde, branco com laranja, nada, nada, laranja, nada, nada:



Este é um cabo fora do padrão, que não deve ser usado em instalações, mas, em compensação, ocupa um volume muito menor e pode ser útil em emergências.



Outro componente que pode ser útil em algumas situações é o conector de loopback, que é usado por programas de diagnóstico para testar a placa de rede. Ele é feito usando um único par de fios, ligado nos contatos 1, 2, 3 e 6 do conector, de forma que os dois pinos usados para enviar dados sejam ligados diretamente nos dois pinos de recepção, fazendo com que a placa receba seus próprios dados de volta:

A pinagem do conector de loopback é:

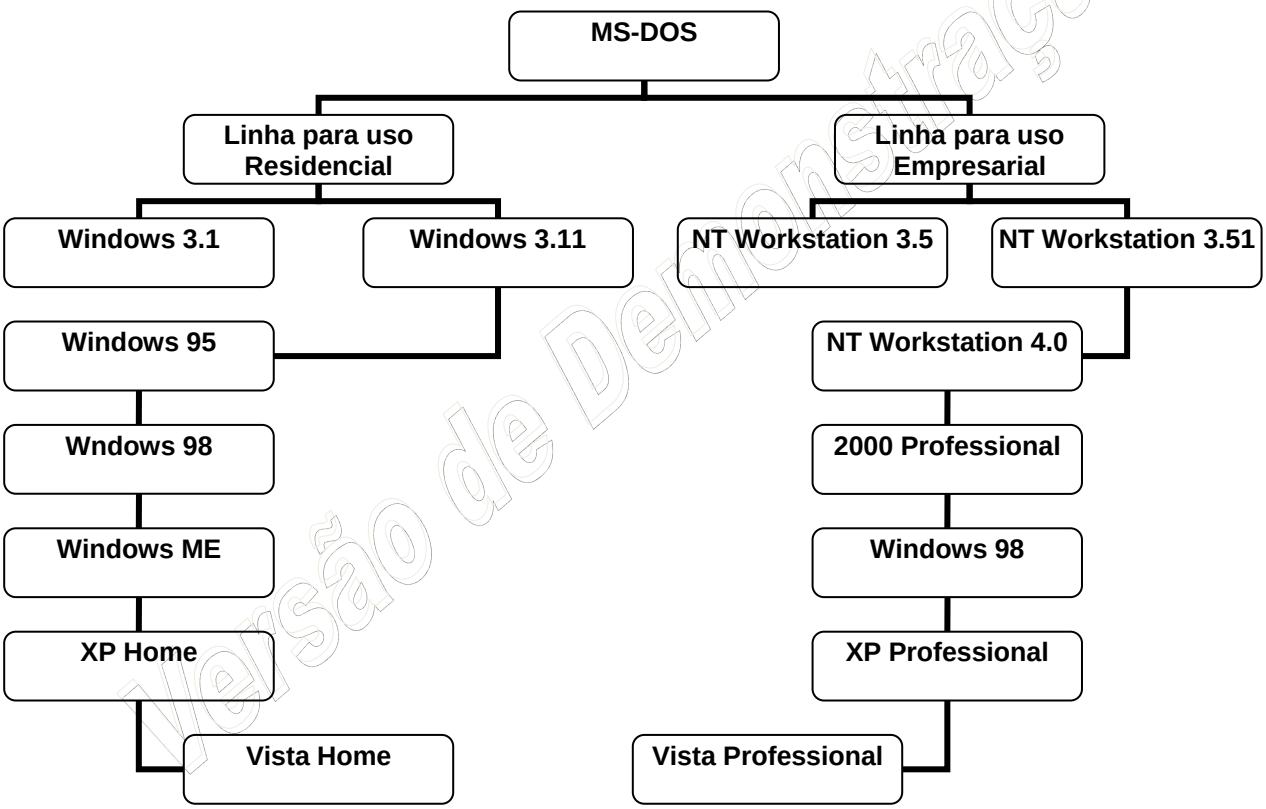
1- Branco com laranja
2- Laranja
3- Branco com laranja (retornando)
4- nada
5- nada
6- Laranja (retornando)
7- nada
8- nada

Ao plugar o conector na placa de rede, você notará que o link da rede é ativado. Ao usar o comando "mii-tool" no Linux, por exemplo, você teria um "eth0: no link" com o cabo de rede desconectado e passaria a ter um "eth0: negotiated 100baseTx-FD, link ok" depois de encaixar o conector de loopback.

11 COMPETÊNCIA 6 – SISTEMAS OPERACIONAIS CLIENTES DE REDE

11.1 FAMILÍLIA DE SISTEMAS OPERACIONAIS CLIENTES

Agora que sabemos como funcionam as redes vamos estudar um pouco sobre os clientes da rede. Em especial, iremos adotar neste curso os sistemas operacionais proprietários da Microsot. Para começar, iremos estudar a evolução dos sistemas operacionais clientes ao longo dos anos, e aprofundaremos nossos estudos com a adoção do Windows XP Professional, por ser este ainda o mais utilizado nos dias de hoje.



A Microsoft inicia suas operações internacionais com o sistema operacional MS-DOS, de Microsoft Data Operation System, por volta de 1986. Sua criação está diretamente relacionada com a explosão de consumo e produção dos IBM/PC. Os primeiros computadores pessoais vendidos massivamente. O MS-DOS era composto apenas por uma interface de console, muito parecida com a interface não-gráfica dos Linux hoje em dia.

Relativamente simples de operar, possuía como núcleo principal de interação com o usuário o programa command.com, localizado na raiz do disco rígido. O command.com era uma coleção de outras ferramentas, como: dir, move, attrib, del, copy, mem, type, etc. Associado com outras ferramentas, como: deltree, format, fdisk, etc.

o

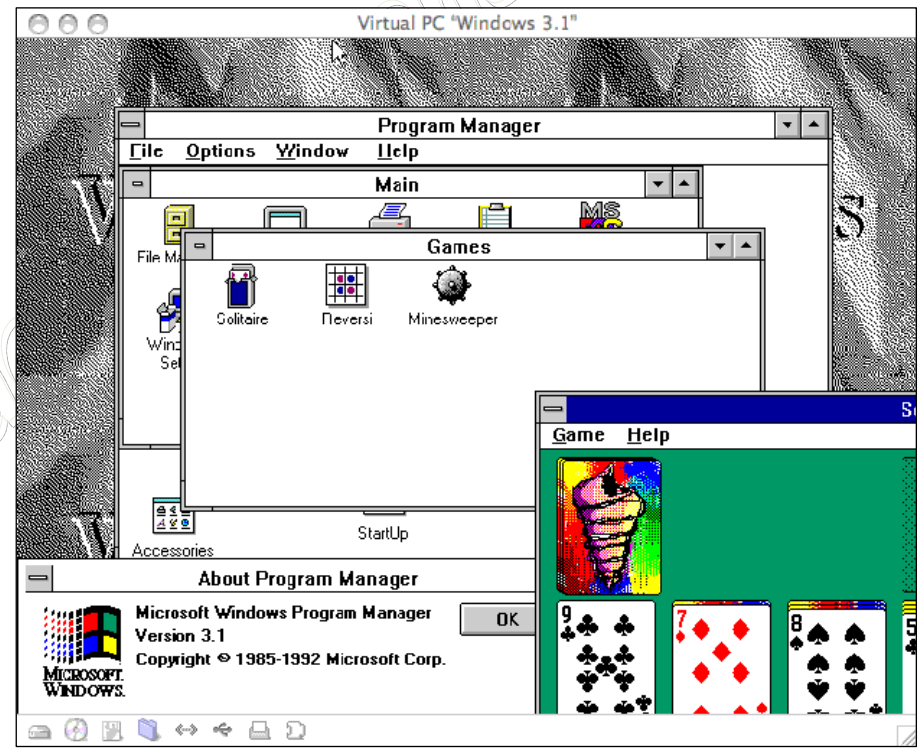
```
C:\>mem

Memory Type        Total   =   Used   +   Free
----------------  --------   --------   --------
Conventional         640K        76K       564K
Upper                 19K         0K        19K
Reserved               0K         0K         0K
Extended (XMS)*    31 661K    2 573K    29 088K
----------------  --------   --------   --------
Total memory       32 320K    2 649K    29 671K

Total under 1 MB      659K        76K       583K

Total Expanded (EMS)               32 000 (32 768 000 bytes
Free Expanded (EMS)*               29 328 (30 031 872 bytes

* EMM386 is using XMS memory to simulate EMS memory as needed.
  Free EMS memory may change as free XMS memory changes.

Largest executable program size       560K (573 520 bytes)
Largest free upper memory block        16K  (16 368 bytes)
MS-DOS is resident in the high memory area.
```

a

o

Como sucessor do MS-DOS, porém ainda dependente do MS-DOS, surge Windows. As versões iniciais do Windows, 1.0 e 2.0, foram pouco utilizadas no Brasil. A primeira versão tornar-se popular de fato foi o Windows 3.1, por volta de 1992. O sistema apresentava uma interface gráfica com ícones para interação com usuário.

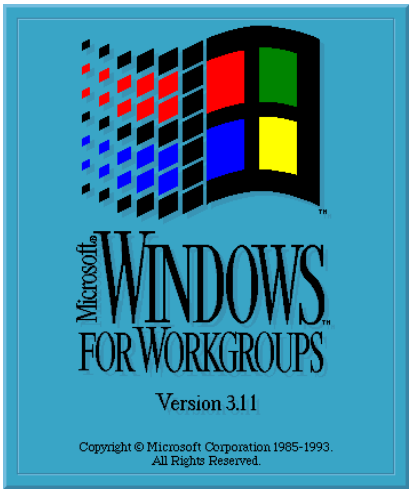
Para aqueles que tiverem a curiosidade de testar versões antigas do MS-DOS ou do Windows, é possível baixar gratuitamente no site da microsoft, http://www.microsoft.com/windows/products/winfamily/virtualpc/default.mspx, o software de máquinas virtuais, o MS Virtual PC.



O site http://www.kernelthread.com/mac/vpc/win.html, apresenta uma relação de sistemas possíveis de serem instalados no MS Virtual PC. Através do Virtual PC é possível instalar e executar outro sistema operacional sem precisar modificar nada em seu atual computador.

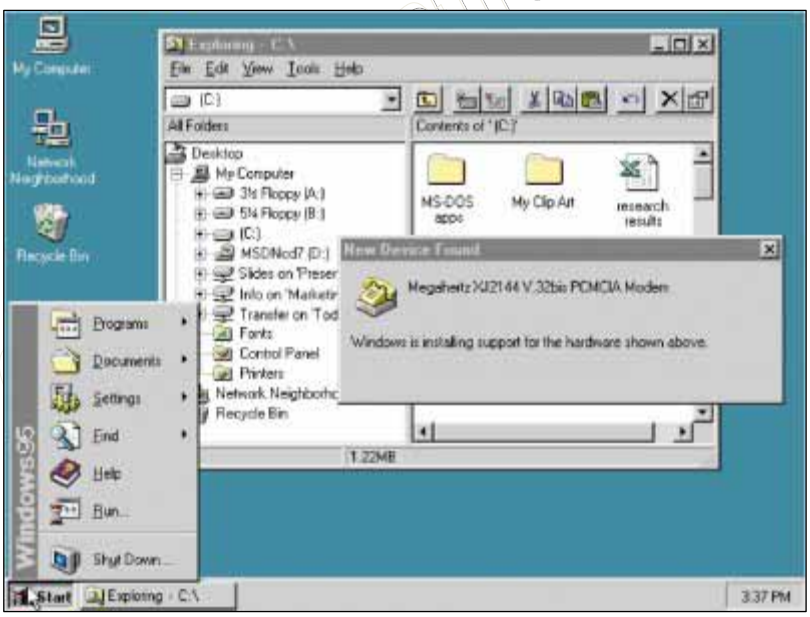
Uma observação importante sobre o Windows 3.1 é que ele não é classificado tecnicamente como um sistema operacional e sim como uma aplicação. Como aplicação ele depende do sistema operacional MS-DOS para poder ser executado. Muitos especialistas classificam o Windows 3.1 como mais do que uma aplicação, como um ambiente operacional, em função de que ele serve de suporte para a execução de várias outras aplicações que não podem ser executadas em MS-DOS nativo.

Por volta de 1993, com a evolução das redes de computadores e consequentemente a necessidade de compartilhamento de periféricos e arquivos, foi lançado o Windows 3.11, também conhecido como Windows for Workgroups.



As diferenças básicas em relação ao Windows 3.1 é que o Windows 3.11 fornecia um suporte melhorado para trabalho em rede e um pouco mais de estabilidade em relação ao Windows 3.1. Esta foi a última versão do Windows baseada na tecnologia de 16 bits.

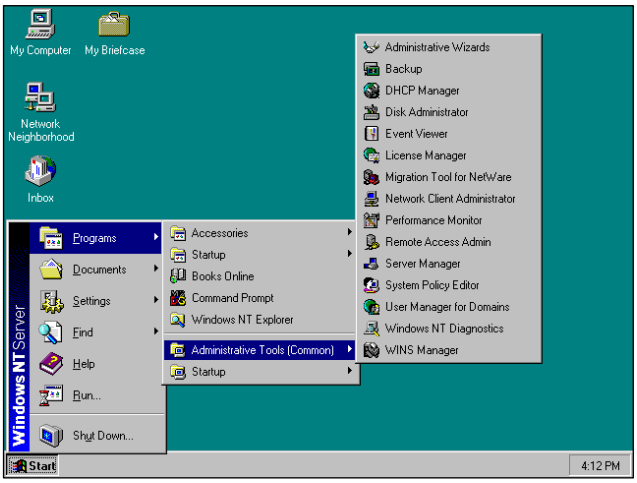
Em 25 de Agosto de 1995 uma nova revolução mudaria os computadores para sempre. Lançado o Windows 95. Um sistema operacional baseado na tecnologia de 32 bits, com uma interface completamente nova em relação às versões anteriores do Windows.



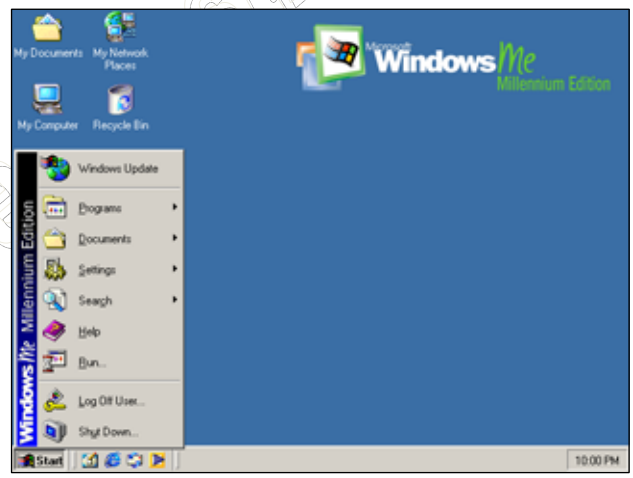
O botão Iniciar, a barra de tarefas, o explorer, entre outros elementos que hoje são muito bem conhecidos, foram novidades trazidas pelo Windows 95. Nesta mesma época a Microsoft já disponibilizava versões do NT Workstation e do NT Server, indicados para uso empresarial das estações de trabalho em rede.

A estratégia da Microsoft em trabalhar com duas linhas de produção, Windows 3.1 ou 95 e NT, geraram confusões e problemas entre os usuários. Por um lado, a Microsoft defendia que a linha empresarial precisava ser mais estável, ou seja, menos susceptível a erros de softwares ou drivers. Para isso compilou o seu kernel com poucas opções de periféricos, aumentou a segurança contra a execução de aplicativos e aumentou o suporte às tecnologias de rede existentes. Por outro lado, isso tornou o NT um sistema de difícil operação, pouco atrativo e de fato voltado para aplicações exclusivamente empresariais. Usuários domésticos que se aventuraram a usar NT acabaram percebendo a necessidade de hardware mais potente, jogos não eram executados, muitos aplicativos legados do Windows 95 não eram mais suportados no NT. Em fim, o NT começou a receber muitas críticas, ora positivas pela estabilidade e segurança, ora negativas pela ausência de suporte a softwares e periféricos, e necessidade de hardware mais potente para ser executado.

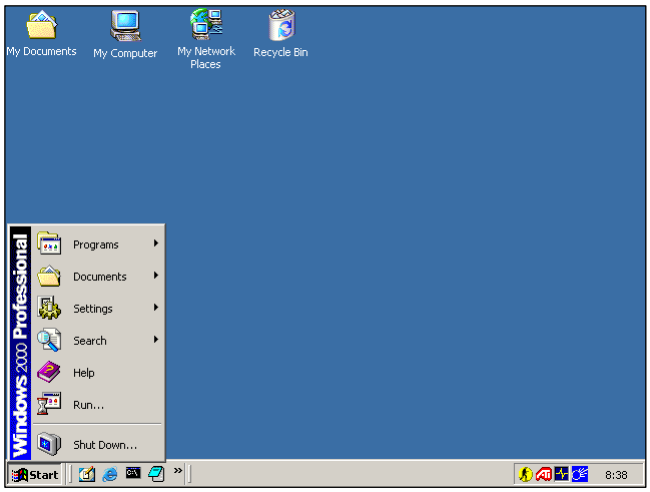
Neste momento a Microsoft já falava em unificar as duas linhas do Windows. Uma nova versão do NT foi lançada: NT Workstation 4.0 e NT Server 4.0. Esta era a versão do NT baseada na tecnologia de 32 bits e com cara de Windows 95. Melhorias substanciais foram feitas em relação a versão anterior do NT. Muitos acreditaram ser esta a versão unificada prometida, tanto que muitas empresas e usuários domésticos começaram a adotar o NT Workstation 4.0 como sistema operacional para as estações da rede e seus computadores pessoais.



Contudo, a robustez do NT persistia, e usuários residenciais começaram a perceber as vantagens que o Windows 95 ainda trazia em relação ao novo NT 4.0, em relação a jogos e periféricos. O golpe final de decisão entre o Windows 95 ou o NT 4.0 aconteceu com o lançamento do Windows 98 e em seguida do Windows ME. O Windows 98 trouxe melhorias significativas em relação ao Windows 95, como estabilidade, segurança e suporte a novos hardwares, contudo sem muitas novas aplicações.



O Windows ME, Millenium Edition, por sua vez, trouxe inovações nos assistentes de instalação, recursos visuais, suporte a novas tecnologias como PnP e USB. Era visualmente mais agradável, porém requisitava um hardware mais robusto. Tornou-se a primeira opção de consumo pelo fato de que o Windows 98, com seus diversos patches de atualizações, já não estava mais sendo suportado em funções de problemas de segurança na Internet e vírus.



Paralelamente ao Windows ME, e mantendo a divisão das linhas de produtos, a Microsoft lança o Windows 2000, nas edições Professional e Server. Embora muitos duvidassem da aceitação do Windows 2000, o fato é que a aceitação deste foi um grande sucesso e muitas empresas adotaram a nova versão.

O objetivo inicial da Microsoft era que o Windows 2000 realizasse o sonho da unificação entre as duas linhas do Windows. Algumas integrações já estavam acontecendo, como por exemplo, um modelo de Drivers para dispositivos de Hardware comum às duas linhas, drivers estes baseados na tecnologia WDM – Windows Driver Model, utilizada tanto no Windows 98 quanto no Windows 2000.

Em 2001 foi lançado o Windows XP. Segundo a Microsoft XP de Experience. O Windows XP, lançado em duas versões: Home e Professional, representa o passo mais importante da Microsoft rumo a unificação das duas linhas do Windows. O XP apresenta uma interface completamente nova, combinando a facilidade do Windows 95/98/Me, com a estabilidade, confiança e segurança do Windows 2000.

Com o desenvolvimento de novas tecnologias para hardwares de servidores, a entrada de novos players no mercado da computação corporativa, e a grande demanda de consumo de todo tipo de empresa sobre a linha corporativa, a Microsoft opta em manter a divisão de sua linha de produtos. Lança em 2003 o Windows Server 2003.

Em 2007 são lançadas as novas versões do Windows para usuários: Vista. E com promessas de lançamento de um novo Windows Server 2008 em 2008. Com estes anúncios a Microsoft oficializa o não interesse em separar as linhas de produtos Windows.

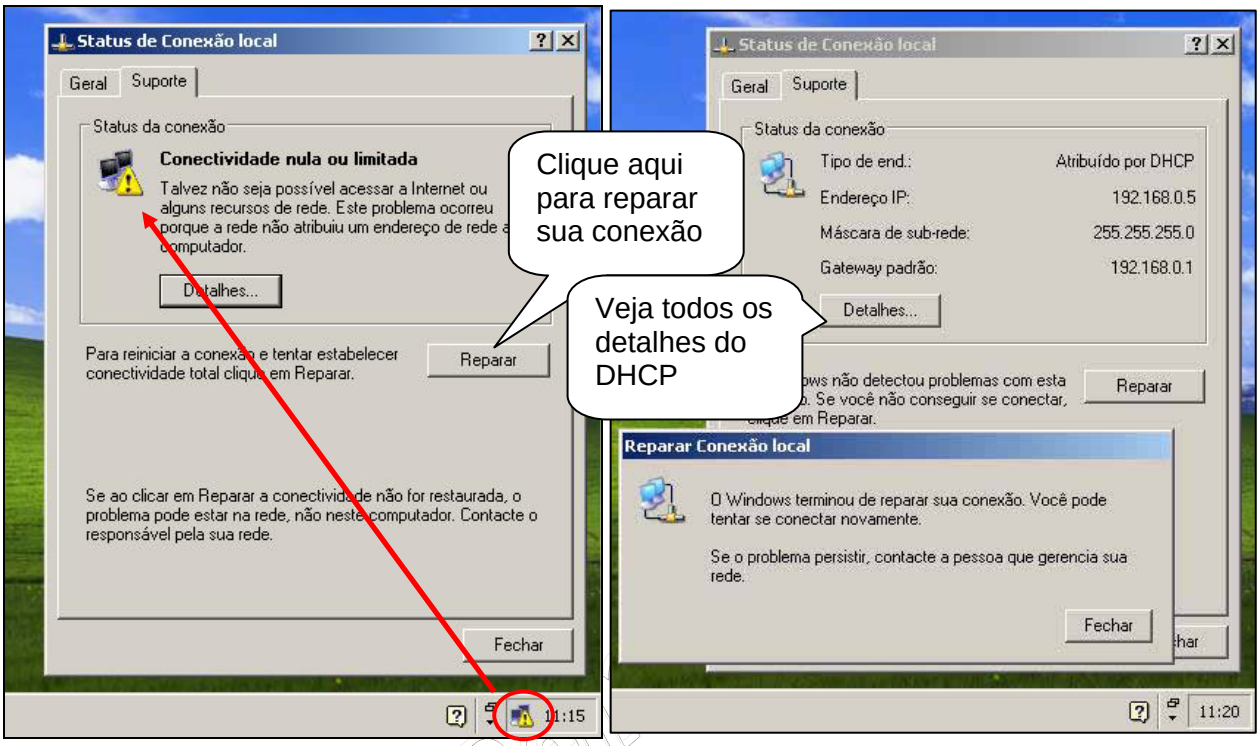
O Windows Vista já é ofertado no mercado Brasileiro através de seis edições:

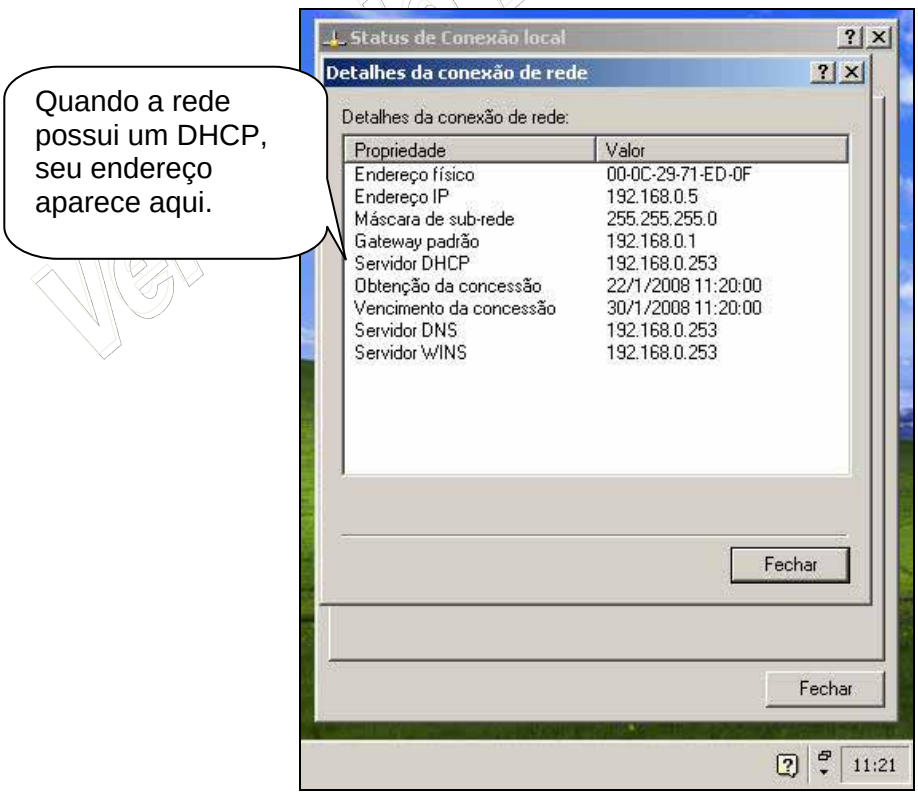
- Vista Starter Edition: versão voltada para usuários não experientes e sem recursos aprimorados como as janelas rotativas em 3D (Aero), é a versão voltada para o público mais carente financeiramente;
- Vista Home Basic: similar ao XP Home Edition, voltada para usuários residenciais, com pacotes extras de aplicativos anti-malwares e recursos avançados de multimídia, porém não provê suporte nem serviços para operações em rede (em especial com Active Directory);
- Vista Home Premium: versão aprimorada da Home Basic, provendo maior suporte para recursos multídias, suporte para HDTV (Televisão Digital de Alta Definição) e o software Windows Media Center, utilizado para controlar o computador através de televisões;
- Vista Business: similar ao XP Professional, versão voltada para empresas de pequeno e médio porte. Conta com serviços e ferramentas de terceiros ou da própria Mirosoft para operações em rede;
- Vista Enterprise: ofertado para as empresas de grande porte, oferece nativamente suporte ao Virtual PC, software de máquinas virtuais; interface com suporte a múltiplos idiomas e a possibilidade de fazer backups ou encriptar grandes volumes de dados;
- Vista Ultimate: a edição mais completa. Tem todas as funcionalidades das versões anteriores e novos serviços online ligados a música, filmes e entreterimento doméstico, incluindo ferramentas para aumentar a performance dos jogos eletrônicos.

Um detalhe especial é que todas as versões do Windows Vista vêm no mesmo DVD de instalação, sendo que a versão a ser instalada depende do CD Key digitado. Será possível atualizar de uma versão a outra, apenas precisando comprar um novo CD Key, que inutilizará o outro. A única exceção ocorre com o Windows Vista Starter Edition, nesta versão, você poderá apenas instalar a nova versão sobre a Starter Edition porém, inutilizando as configurações e programas instalados anteriormente. Existe uma versão em CD do Vista Starter que não possui as outras versões.

11.2 CONFIGURAÇÃO DO TCP/IP

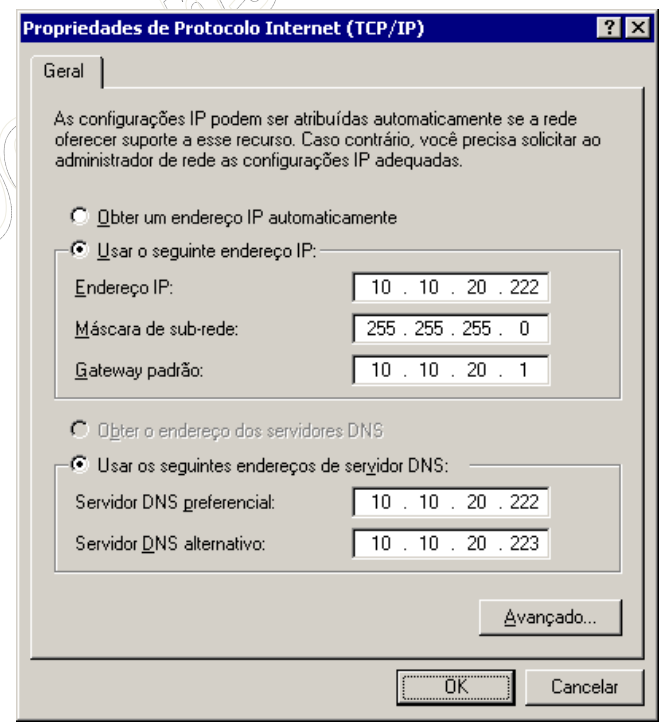
A primeira ação que realizamos em rede é ligar a estação de trabalho ao mesmo switch onde está o servidor. Quando não temos um servidor DHCP na rede, ou não estamos conectados no mesmo barramento de rede que o servidor DHCP, então o sistema apresenta uma mensagem de “Conectividade nula ou limitada”, porém, quando estivermos ao alcance do servidor DHCP, teremos uma séria de informações sobre a rede preenchidas de forma automática.


Status de Conexão local
Geral
Suporte
Status da conexão
Conectividade nula ou limitada
Talvez não seja possível acessar a Internet ou alguns recursos de rede. Este problema ocorreu porque a rede não atribuiu um endereço de rede a este computador.
Detalhes...
Para reiniciar a conexão e tentar estabelecer conectividade total clique em Reparar.
Reparar
Se ao clicar em Reparar a conectividade não for restaurada, o problema pode estar na rede, não neste computador. Contacte o responsável pela sua rede.
Fechar
11:15
Clique aqui para reparar sua conexão
Veja todos os detalhes do DHCP
Tipo de end.: Atribuído por DHCP
Endereço IP: 192.168.0.5
Máscara de sub-rede: 255.255.255.0
Gateway padrão: 192.168.0.1
Reparar Conexão local
O Windows terminou de reparar sua conexão. Você pode tentar se conectar novamente.
Se o problema persistir, contacte a pessoa que gerencia sua rede.
11:20

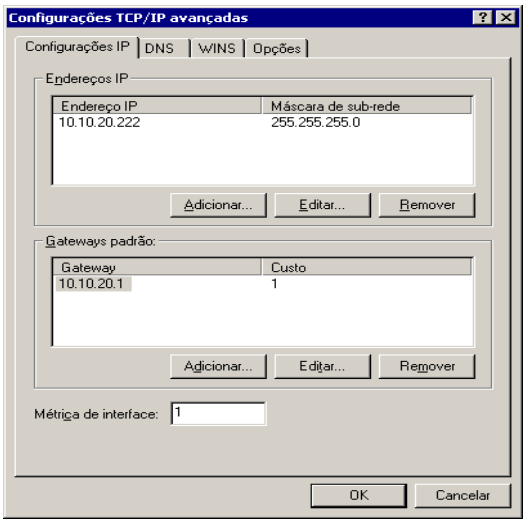

Quando a rede possui um DHCP, seu endereço aparece aqui.
Status de Conexão local
Detalhes da conexão de rede
Detalhes da conexão de rede:
Propriedade | Valor
Endereço físico | 00-0C-29-71-ED-0F
Endereço IP | 192.168.0.5
Máscara de sub-rede | 255.255.255.0
Gateway padrão | 192.168.0.1
Servidor DHCP | 192.168.0.253
Obtenção da concessão | 22/1/2008 11:20:00
Vencimento da concessão | 30/1/2008 11:20:00
Servidor DNS | 192.168.0.253
Servidor WINS | 192.168.0.253
Fechar
Fechar
11:21

Entretanto, se sua rede não possui um servidor DHCP, então você deverá realizar a configuração do TCP/IP manualmente. Para acessar as propriedades da interface de rede a ser configurada, siga os passos indicados a seguir:

1. Faça o logon como Administrador ou com uma conta com permissão de administrador.
2. Abra o Painel de controle: Iniciar -> Configurações -> Painel de controle.
3. Dê um clique duplo na opção Conexões dial-up e de rede.
4. Será exibida uma janela com todas as conexões disponíveis. Clique com o botão direito do mouse na conexão a ser configurada e, no menu de opções que é exibido, clique em Propriedades.
5. Pronto, será exibida a janela de propriedades da conexão, na qual você poderá fazer diversas configurações.
6. Na janela de propriedades da conexão dê um clique em Protocolo Internet (TCP/IP) para selecioná-lo.
7. Clique em Propriedades. Nesta janela você deve informar se as configurações do TCP/IP serão obtidas a partir de um servidor DHCP (Obter um endereço IP automaticamente) ou se estas configurações serão informadas manualmente (Usar o seguinte endereço IP). Ao marcar a opção Usar o seguinte endereço IP, você deverá informar um número IP a ser utilizado, a máscara de sub-rede, o número IP do Gateway padrão e o número IP de um ou dois servidores DNS, conforme exemplo da Figura a seguir:



8. Além das configurações básicas, da tela da Figura anterior, você pode configurar uma série de opções avançadas do protocolo TCP/IP. Para acessar a janela de configurações avançadas, clique em Avançado... Será aberta a janela de configurações avançadas, com a guia Configurações IP selecionada por padrão, conforme indicado na Figura a seguir:

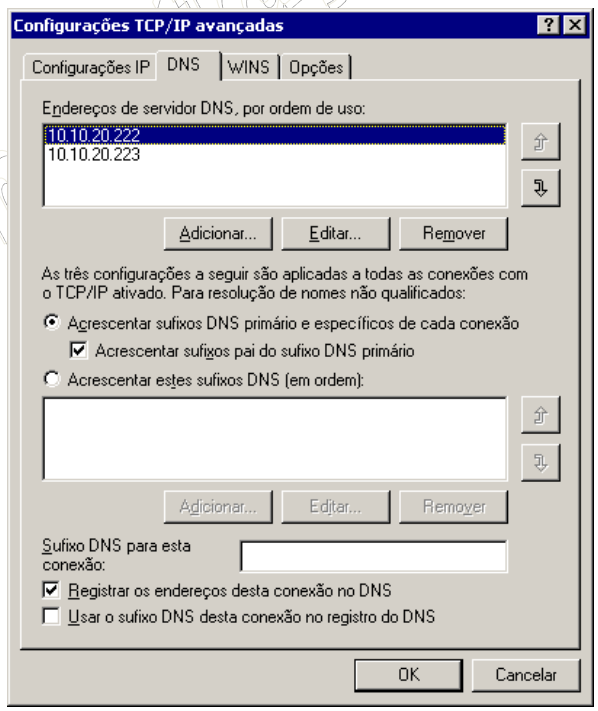


9. É possível ter mais de um endereço IP associado com a mesma placa de rede. O que não é permitido é ter o mesmo número IP, associado a duas ou mais placas de rede. Para adicionar um novo número IP, clique em **Adicionar...,** abaixo da lista de endereços IP configurados. Será aberta a janela Endereço TCP/IP (muito mal traduzida por sinal). Para adicionar um novo endereço basta digitá-lo no campo IP, digite a respectiva máscara de sub-rede e clique em Adicionar. Você estará de volta à janela de configurações avançadas do TCP/IP e o novo endereço IP já será exibido na

lista. A partir de agora, a nova interface está com dois endereços IP. Você pode adicionar mais endereços IP, utilizando o botão Adicionar... e preenchendo as informações necessárias.

10. Você também pode ter mais de um default gateway configurado. Neste caso, quando o primeiro default gateway da lista estiver indisponível, o TCP/IP tenta utilizar o segundo e assim por diante. Para adicionar mais um default gateway, clique em Adicionar..., abaixo da lista de default gateways configurados. Será aberta a janela para que você informo o número IP do novo default gateway e o respectivo custo, em número de hopes. Se você quer que um default gateway seja utilizado somente como contingência, no caso de nenhum outro gateway estar disponível, configure-o com um valor elevado para o custo. Digite as informações do novo gateway e clique em OK. Pronto, o novo número já será exibido na guia de Configurações IP.

11. Clique na guia DNS. Serão exibidas as opções indicadas na Figura a seguir:



Nesta guia você informa o endereço IP de um ou mais servidores DNS. Para acrescentar novos servidores, basta utilizar o botão Adicionar... Você pode alterar a ordem dos servidores DNS na lista, clicando nos botões com o desenho de uma flecha para cima ou para baixo. É importante descrever como o Windows utiliza a lista de servidores DNS. As consultas são enviadas para o primeiro servidor da lista. Se este servidor não conseguir responder a consulta, esta não será enviada para os demais servidores da lista. O segundo servidor da lista somente será pesquisado se o primeiro servidor estiver off-line e não estiver respondendo; o terceiro servidor da lista somente será pesquisado se o primeiro e o segundo servidor DNS estiverem off-line e não estiverem respondendo e assim por diante. Nesta guia você também pode configurar as seguintes opções:

- **Acrescentar sufixo DNS primário e específicos de cada conexão**: O sufixo DNS é configurado na guia Identificação de rede, das propriedades do meu Computador. Por exemplo, um computador com o nome micro01.abc.com, tem como sufixo DNS abc.com. Esta opção especifica que a resolução de nomes não qualificados (por exemplo micro01.abc.com é um FQDN, ou seja, um nome completamente qualificado, já micro01 é um nome não qualificado, ou seja, sem o domínio como sufixo) usados neste computador seja limitada aos sufixos do domínio do sufixo primário e todos os sufixos específicos da conexão. Os sufixos específicos da conexão são configurados em Sufixo DNS para esta conexão. O sufixo DNS primário é configurado clicando em Propriedades, na guia Identificação de rede (disponível em Sistema, no Painel de controle). Por exemplo, se o sufixo do seu domínio primário for abc.com e você digitar ping xyz em um prompt de comando, o Windows 2000 consultará xyz.abc.com. Se você também configurar um nome de domínio específico de conexão em uma das suas conexões para vendas.abc.com, o Windows 2000

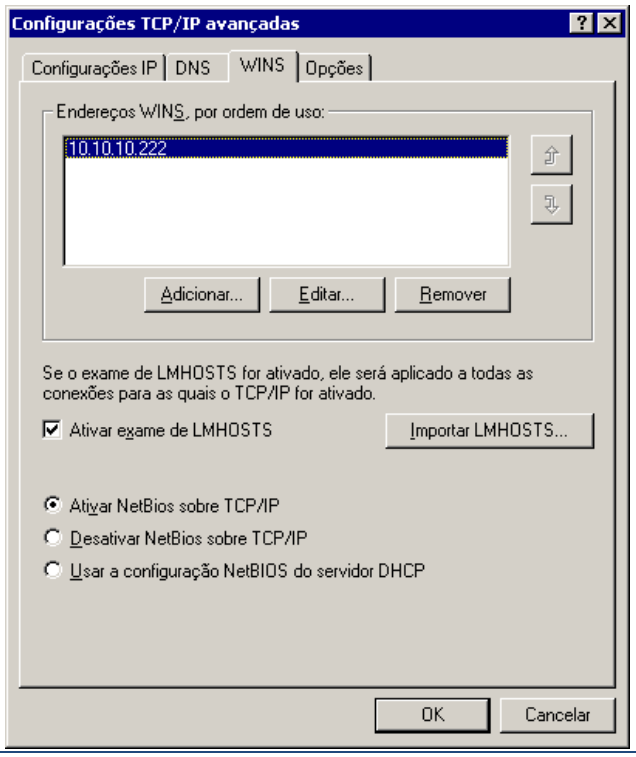
consultará xyz.abc.com e xyz.vendas.abc.com. A lista de domínios que será pesquisada, quando você digita um nome não qualificado, também é definida nesta guia, conforme será explicado logo a seguir.

- **Acrescentar sufixos pai do sufixo DNS primário:** Especifica se a resolução de nomes não qualificados usados neste computador inclui os sufixos pai do sufixo DNS primário e o domínio de segundo nível. O sufixo DNS primário é configurado clicando em Propriedades na guia Identificação de rede (disponível na opção Sistema do Painel de controle). Por exemplo, se o sufixo DNS primário for vendas.abc.com e você digitar ping xyz no prompt de comando, o Windows 2000 também consultará vendas.abc.com e abc.com.

- **Acrescentar estes sufixos DNS (em ordem):** Especifica que a resolução de nomes não qualificados usados neste computador seja limitada aos sufixos do domínio listados no campo "Acrescentar estes sufixos DNS". Os sufixos DNS específicos da conexão e primários não serão usados para resolução de nomes não qualificados. Ao marcar esta opção, você deve especificar uma lista de sufixos que deverá ser utilizada, para a tentativa de resolução de nomes não qualificados. Por exemplo, se nesta lista você acrescentar os seguintes sufixos: sul.vendas.abc.com, vendas.abc.com e abc.com, nesta ordem, ao digitar ping xyz, o Windows tentará localizar este host, utilizando os seguintes nomes: xyz.sul.vendas.abc.com, xyz.vendas.abc.com e xyz.abc.com. Para acrescentar um novo sufixo basta marcar esta opção e utilizar o botão Adicionar. Você também pode alterar a ordem dos sufixos clicando nos botões com a seta para cima e seta para baixo. Para remover um sufixo basta selecioná-lo na lista e clicar em Remover.

- **Registrar endereços desta conexão no DNS:** Especifica que o computador tente o registro dinâmico no DNS, dos endereços IP desta conexão com o nome completo deste computador, como especificado na guia Identificação de rede (disponível em Sistema no Painel de Controle).

- **Usar o sufixo DNS desta conexão no registro do DNS:** Especifica se a atualização dinâmica do DNS será usada para registrar os endereços IP e o nome de domínio específico desta conexão. O nome DNS específico desta conexão é a concatenação do nome do computador (que é o primeiro rótulo do nome completo do computador) e o sufixo DNS desta conexão. O nome completo do computador é especificado na guia Identificação de rede (disponível em Sistema, no Painel de controle). Se a caixa de seleção Registrar os endereços desta conexão no DNS estiver selecionada, o registro é uma adição ao registro do DNS do nome completo do computador.



12. Defina as configurações desejadas e clique na guia WINS. Serão exibidas as opções indicadas na Figura a seguir:

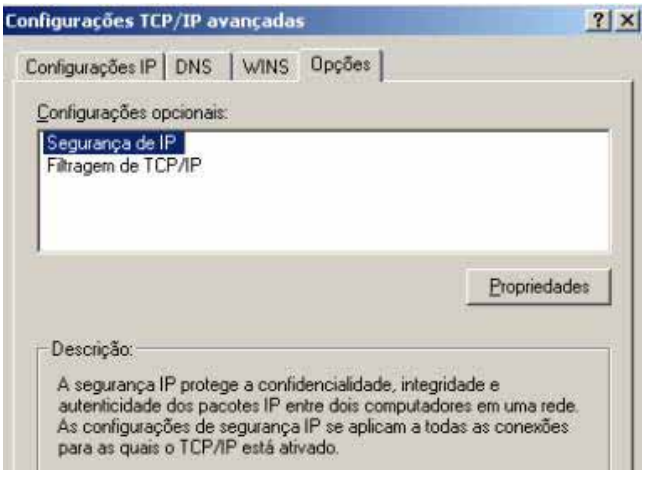
13. Nesta guia você informa o endereço IP de um ou mais servidores WINS. Para acrescentar novos servidores, basta utilizar o botão Adicionar... Você pode alterar a ordem dos servidores WINS na lista, clicando nos botões

com o desenho de uma flecha para cima ou para baixo. É importante descrever como o Windows utiliza a lista de servidores WINS. As consultas são enviadas para o primeiro servidor da lista. Se este servidor não conseguir responder a consulta, esta não será enviada para os demais servidores da lista. O segundo servidor da lista somente será pesquisado se o primeiro servidor estiver off-line e não estiver respondendo; o terceiro servidor da lista somente será pesquisado se o primeiro e o segundo servidor WINS estiverem off-line e não estiverem respondendo e assim por diante. Nesta guia você também pode configurar as seguintes opções:

- **Ativar exame de LMHOSTS:** Especifica se será usado um arquivo Lmhosts para a resolução de nomes NetBIOS. O arquivo Lmhosts será usado para resolver os nomes de NetBIOS de computadores remotos para um endereço IP. Clique em Importar LMHOSTS para importar um arquivo para o arquivo Lmhosts.
- **Ativar NetBios sobre TCP/IP:** Especifica que esta conexão de rede usa o NetBIOS sobre TCP/IP (NetBT) e o WINS. Quando um endereço IP é configurado manualmente, esta opção é selecionada por padrão para ativar o NetBIOS e o uso do WINS para este computador. Essa configuração será necessária se este computador se comunicar pelo nome com computadores que usam versões anteriores do Windows (Windows 95/98, NT 4.0, etc.). Antes de alterar esta opção, verifique se não é necessário usar nomes de NetBIOS para esta conexão de rede. Por exemplo, se você se comunicar somente com outros computadores que estejam executando o Windows 2000 ou computadores na Internet que usam o DNS.
- **Desativar NetBios sobre TCP/IP:** Desativa o uso de NetBios sobre TCP/IP. Pode ser utilizada em uma rede baseada apenas em versões do Windows tais como Windows 2000, Windows XP e Windows Server 2003.
- **Usar a configuração NetBios do servidor DHCP:** Especifica que esta conexão de rede obtenha suas configurações de NetBIOS sobre TCP/IP (NetBT) e de WINS, a partir de um servidor DHCP.

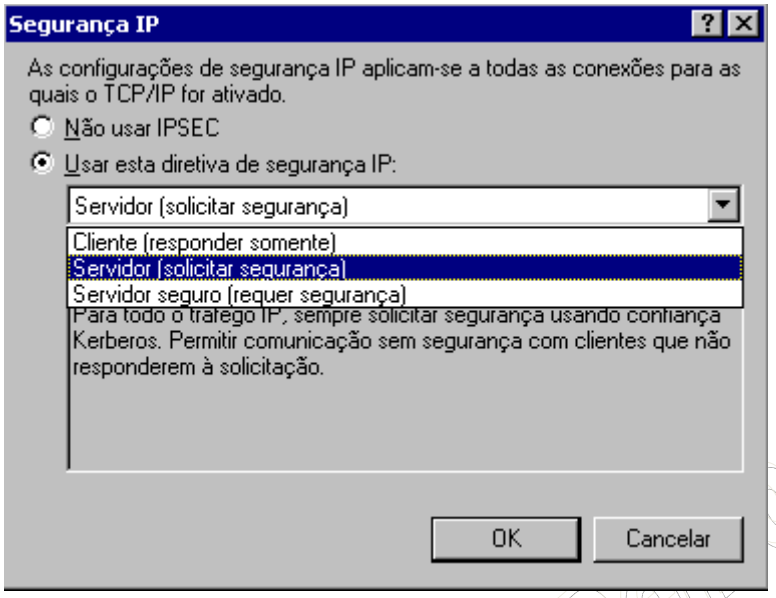
Quando um endereço IP é obtido automaticamente, esta opção fica selecionada por padrão de forma que o computador use as definições de configuração do NetBT conforme elas forem sendo fornecidas opcionalmente pelo servidor DHCP quando ele obtiver um endereço IP usando o DHCP. Você deve selecionar esta opção somente se o servidor DHCP estiver configurado para fornecer todas as opções de configuração de WINS para os clientes.

14. Defina as configurações desejadas e clique na guia Opções. Serão exibidas as opções indicadas na Figura a seguir:



15. Nesta janela você pode configurar se a interface que está sendo configurada deve ou não utilizar uma das diretivas de IPSec habilitadas (caso haja alguma diretiva habilitada) e também pode definir filtros com base no protocolo e na porta de comunicação. Para habilitar o uso de uma das diretivas do IPSec, clique em Segurança de IP para marcar esta opção e em seguida clique em Propriedades.

16. Será aberta a janela Segurança de IP. Para habilitar o IPSec clique em Usar esta diretiva de segurança IP e, na lista de diretivas, selecione a diretiva a ser aplicada, conforme exemplo da Figura a seguir e clique em OK. Você estará e volta à janela de propriedades Avançadas do TCP/IP.

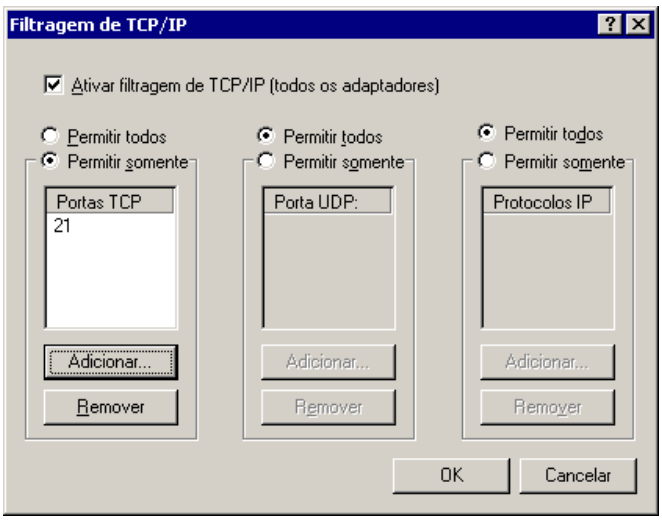


17. Para definir um filtro clique em Filtragem de TCP/IP e em seguida no botão Propriedades. Será exibida a janela para definição de filtros. Nesta janela você tem as seguintes opções:

- **Ativar filtragem de TCP/IP (todos os adaptadores):** Ao marcar esta opção você especifica se a filtragem de TCP/IP será ativada para todos os adaptadores. A filtragem de TCP/IP especifica os tipos de tráfego de entrada destinados para este computador que serão permitidos. Para configurar a filtragem de TCP/IP, selecione esta caixa de seleção e especifique os tipos de tráfego TCP/IP permitidos para todos os adaptadores neste computador em termos de protocolos IP, portas TCP e portas UDP. Você deve ter cuidado ao usar os filtros, para não desabilitar portas que sejam necessárias para os serviços básicos de rede, tais como DNS, DHCP, compartilhamento de pastas e impressoras e assim por diante.

18. Vamos aplicar um exemplo de filtro. O FTP usa o protocolo TCP na porta 21. Para o nosso exemplo, para as portas TCP, vamos permitir apenas o uso do FTP na porta 21. Marque a opção Ativar filtragem de TCP/IP (todos os adaptadores). Em seguida marque a opção Permitir somente nas portas TCP. Clique em Adicionar... Será exibida a janela adicionar filtro, para que você adicione o número da porta. Digite 21, conforme indicado na Figura a seguir e clique em OK.

19. Você estará de volta à janela Filtragem de TCP/IP, com a porta TCP/21 já adicionada, conforme indicado na Figura a seguir:



20. Clique em OK. Você estará de volta a janela de configurações avançadas do TCP/IP. Clique em OK para fechá-la.

21. Você estará de volta à janela de configurações da interface de rede. Clique em Fechar para fechá-la.

11.3 TESTES DE CONEXÕES PONTO-A-PONTO

O teste de conexão de rede consiste em utilizar o protocolo ICMP para diagnosticar o estado da rede. Este protocolo por sua vez é representado através da ferramenta PING.

Testar uma rede significa pingar outros hosts e servidores do mesmo segmento ou não de rede. Como exemplo. Supondo ser o IP da nossa estação de trabalho o 10.10.20.222, e sendo o endereço IP do Gateway o 10.10.20.1. Diagnosticar o estado da conexão de rede neste caso significaria realizar um ping do IP 10.10.20.222 para o IP 10.10.20.1. Para isso, inicie uma janela do Promtp de Comando e digite:

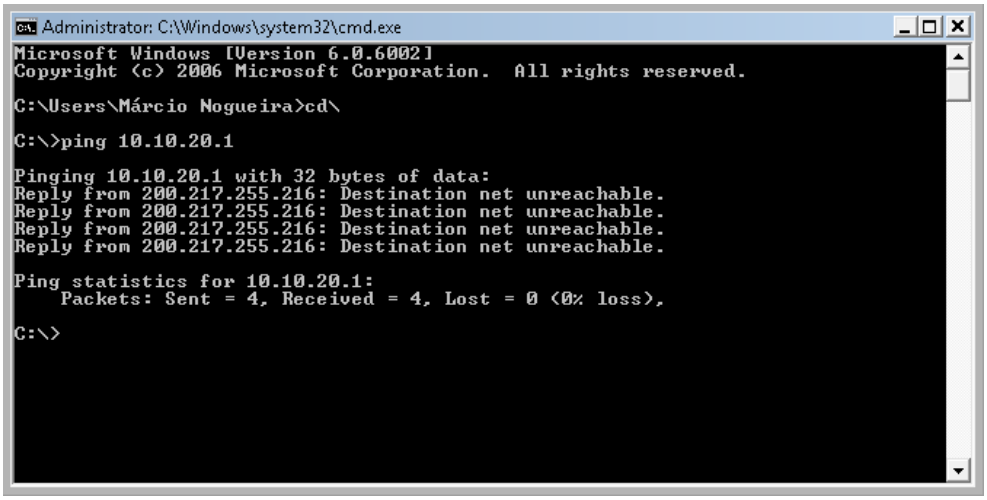
c:\ping 10.10.20.1

Antes de analisarmos o resultado do teste de diagnóstico, convém ordenar uma linha racional e coerente de testes de diagnósticos, onde propomos a seguinte abordagem:

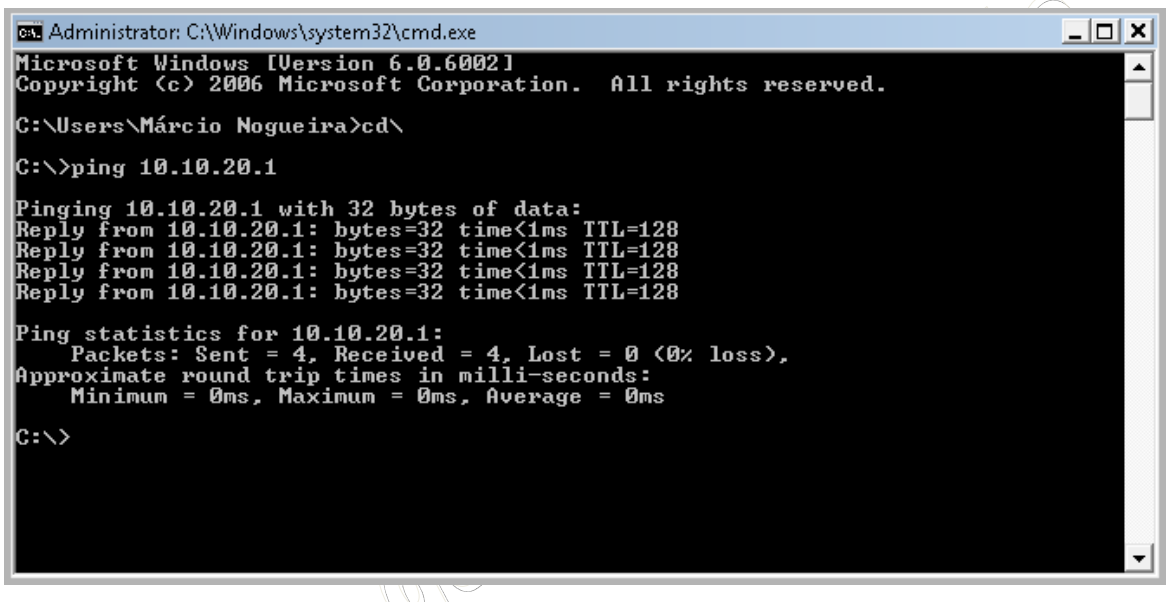
1. Testar o IP 127.0.0.1 (localhost) a fim de identificar se o protocolo TCP/IP está devidamente instalado no computador (caso o ping falhe para este IP, reinstale o protocolo TCP/IP);
2. Testar o próprio IP do computador, neste caso seria o IP 10.10.20.222, a fim de identificar novamente problemas com o protocolo TCP/IP da própria máquina;
3. Testar o IP do servidor (Servidor DHCP, se este estiver presente, ou Servidor Gateway), caso sua rede possua um servidor DHCP e você não esteja recebendo IP deste servidor, logo você não conseguirá pingar o servidor, precisando verificar as questões do cabeamento da rede. Em relação ao gateway, caso este seja um firewall, é bem provável que você também não consiga pingar para ele, neste caso você precisará identificar outra estação de trabalho na rede, que esteja ao alcance do seu segmento de rede, e que também não tenha um firewall que impeça as respostas ICMP do ping.

Agora que temos uma abordagem de diagnóstico de rede, vamos compreender os resultados esperados:

No primeiro caso, vamos supor que sua estação de trabalho não esteja no mesmo segmento de rede do seu destino, neste caso o resultado do comando ping será uma mensagem informando que a rede de destino não está ao alcance. Este tipo de erro geralmente ocorre quando você utiliza uma máscara de rede equivocada ou existe algum problema de roteamento de sub-redes.



No segundo caso, vamos supor que o computador de endereço 10.10.20.222 conseguiu realizar com sucesso o ping para o IP 10.10.20.1. Neste caso, algumas informações serão exibidas, como: a quantidade de pacotes enviadas (no Windows XP o padrão são o envio de 4 pacotes), a quantidade de pacotes respondidas, um percentual de pacotes perdidos, o tamanho em bytes dos pacotes transmitidos (no Windows XP o padrão é o tamanho de 32 bytes, podendo varia de 0 bytes até 65500 bytes) e o tempo de resposta que um pacote levou para ir e voltar do destino.



Cada um desses campos apresenta seu grau de importância, por exemplo:

- O percentual de pacotes perdidos, quando diferente de zero, indica o grau de problemas, podendo ser: super processamento no próprio computador (podendo ser alguma aplicação que esteja consumindo o processador e memória RAM do computador, ou mesmo algum problema de hardware no computador), ou um congestionamento entre o computador e o destino, ou mesmo algum problema de processamento no próprio computador de destino;
- O tempo de resposta (time), por sua vez, nos ajuda a interpretar o percentual de pacotes perdidos ou compreender problemas de congestionamento da rede: quando o problema é de processamento local, então o tempo de resposta é bem pequeno, porém a quantidade de pacotes perdidos é diferente de zero. Quando há algum problema de processamento no destino teremos o tempo de resposta alto, perdas sucessivas de pacotes e mensagens de retorno do tipo “o tempo máximo de resposta foi excedido”. Quando o problema é de congestionamento na rede, o diagnóstico é feito através da comparação do retorno do time com a seguinte tabela de qualidade de redes (abordagem nossa):
    - Redes locais: de 0ms até 1ms;
    - Redes locais com até 05 switchs intermediários: de 1ms até 2ms;
    - Redes locais segmentadas por roteador: de 1ms até 3ms;
    - Extranets interligadas através de conexão dedicada: de 5ms até 10ms;
    - Conexões Dial-UP (discadas): de 100ms até 500ms;
    - Conexões Banda Larga (ADSL, Cable, Satélite): de 50ms até 200ms;

11.4 COMPARTILHAMENTO DE RECURSOS EM REDE

Analisaremos agora as principais ferramentas utilizadas para acesso as redes, focaremos o uso sobre o Microsoft Windows XP Professional, em função deste ainda ser o sistema em maior uso na atualidade e provavelmente, em função do atual mercado de hardwares, a versão a ser predominante nos próximos 3 anos nos ambientes corporativos em rede.

Antes de começarmos, é importante ressaltar que estações de trabalho em redes não estão limitadas ao uso por apenas um único funcionário. É muito comum que durante o dia, mais de um funcionário utilize a mesma estação de trabalho, como no caso de empresas que possuem dois turnos para uma mesma função em departamento. Um funcionário utiliza a máquina pela manhã e, a tarde, um segundo funcionário a utiliza. Isso torna os recursos de acesso a rede mais complexos, pois precisam proteger tanto os dados do usuário 1 quanto do usuário 2, ao mesmo tempo em que libera ou restringe o acesso as recursos compartilhados em rede por usuário, e não por estação de trabalho.

As ferramentas que analisaremos são as ferramentas que não existem nas versões para usuário doméstico do Windows, e que existem tanto no XP quanto no Vista, e que são utilizadas para operações em estações de trabalho em rede:

- Controle de permissão de pastas e arquivos em partições NTFS;
- Remote Desktop;
- Diretivas de Segurança Local;

Controle de permissão de pastas e arquivos em partições NTFS

A partir do Windows NT/2000 um novo recurso está disponível: compartilhar recursos (pastas, arquivos) localmente ou na rede com maior segurança. É o controle de acesso, que permite selecionar quais usuários terão permissão para acessar o objeto compartilhado. O controle de acesso é feito através de permissões: NTFS e de compartilhamento.

As permissões NTFS são válidas tanto localmente (no próprio PC) quanto para a rede : quando um usuário fizer logon, seja no mesmo PC ou em outro qualquer, ele só poderá acessar o recurso compartilhado se tiver permissões adequadas.

As permissões de compartilhamento só têm efeito ao acessar recursos compartilhados na rede, mas não no próprio PC.

As permissões NTFS permitem atribuir permissões a pastas e arquivos, conferindo um alto grau de segurança e controle de acesso a nível de usuário. Têm efeito localmente e através da rede. As permissões definem o tipo de acesso concedido a um usuário ou a um grupo para um objeto: arquivos e pastas, chaves do registro, serviços, impressoras.

É um mecanismo de segurança que determina quais usuários ou grupos estão autorizados a executar quais operações em um objeto.

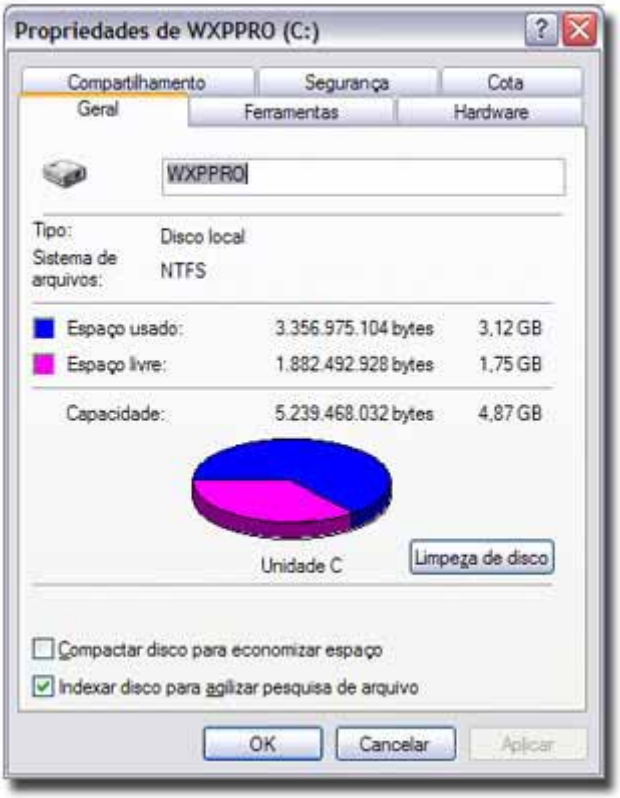
Você pode permitir ou negar acesso aos recursos compartilhados. Você também pode permitir somente permissões adequadas a um usuário: controle total, modificar, ler e executar, listar conteúdo de pastas, ler e gravar.

Como exemplo podemos citar:

1. É possível atribuir permissões a uma impressora compartilhada na rede ou localmente: somente usuários autorizados poderão imprimir, gerenciar e outras pessoas poderão ter suas permissões negadas e não poderão imprimir documentos;
2. É possível atribuir permissões a pastas e arquivos no mesmo PC: seus documentos só poderão ser visualizados a usuários com a permissão de leitura, delegados por você, e ninguém poderá alterá-los, exceto você.

Como pré-requisito para operação dos controles de permissões é necessário: que a unidade que contém as pastas e os arquivos devem estar formatadas em NTFS (WinNT e WinXP) ou NTFS 5 (Win2000 Server) e rodando um sistema operacional Win2000 ou superior.

Para confirmar o tipo de sistema de arquivos formatado em uma unidade de disco: No Windows Explorer, selecione a unidade e clique com o botão direito / Propriedades. Na guia Geral veja Sistema de arquivos.



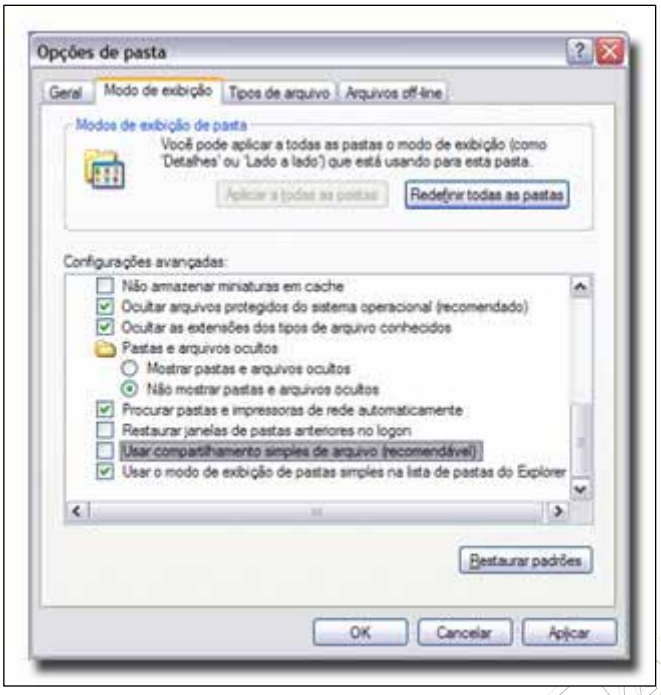
A partição formatada em NTFS tem outras vantagens em relação a FAT32 : segurança (definição de cotas de disco, auditoria de objetos, criptografia, journaling e controle de acesso), espaço livre (compactação de dados), suporte a arquivos com mais de 4 Gb e desempenho.

A única desvantagem é a compatibilidade : o sistema de arquivos NTFS só é reconhecida pelo WinNT, Win2000 e WinXP, enquanto o sistema de arquivos FAT32 é compatível com todos os Windows, exceto WinNT, e versões mais antigas do MS-DOS. Veja mais nessa matéria FAT32 X NTFS.

Podemos falar agora de nível de acesso, como sendo o controle de acesso de usuários e grupos a determinados objetos. Exemplo:

- Um usuário pode ter acesso ao conteúdo de um arquivo, outro fazer alterações, e um outro grupo nem poderá acessar o arquivo do mesmo PC.

Para Alterar permissões de pastas: No Windows Explorer clique com o botão direito no nome da pasta e clique em Compartilhamento e segurança... e na guia Segurança. Se você for membro de um grupo de trabalho e deseja visualizar a guia Segurança, abra o Painel de Controle e clique em Opções de pasta. Na guia Modo de exibição > Configurações Avançadas, desmarque Usar compartilhamento simples de arquivo (recomendável).



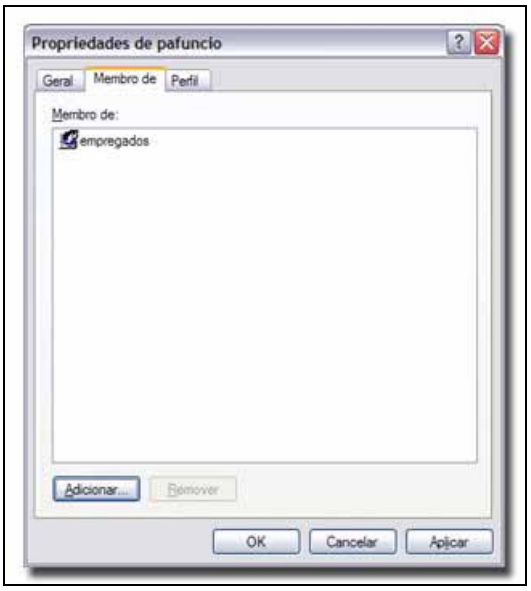
Algumas observações sobre as propriedades de objetos:

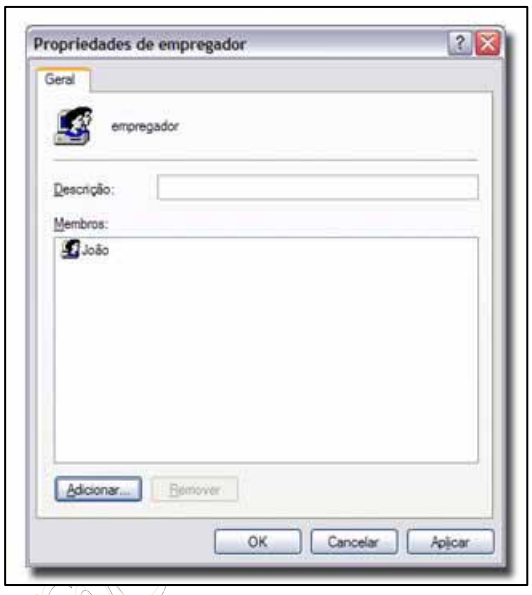
- Proprietários: Todos os objetos têm um proprietário, que por padrão é o criador do objeto. O proprietário poderá sempre alterar as permissões, independentemente das permissões definidas ao objeto;
- Herança: As permissões são automaticamente herdadas do objeto pai. Exemplo : uma subpasta herda as mesmas permissões da pasta que está contida; os arquivos criados dentro de uma pasta herdam as permissões da pasta. Esse recurso permite gerenciar e atribuir permissões com agilidade e facilidade;
- Operações: Ao copiar ou mover um objeto de uma partição para outra, as permissões serão perdidas e as novas serão herdadas do objeto pai. Se a operação for na mesma partição, as permissões serão mantidas. Ao copiar um objeto compartilhado (com permissões) de uma partição NTFS para outra FAT ou FAT32, as permissões serão perdidas;
- Tipo de objetos: As permissões são diferentes dependendo do tipo de objeto (pastas, arquivos ...). As permissões de ler, modificar permissões, alterar proprietário e excluir são comuns a todos os objetos;

Para realizar o compartilhamento de pastas execute os passos a seguir:

1) Faça logon como Administrador (ou usuário que tenha permissões de administrador) onde se encontra a pasta a ser compartilhada e serão definidas as permissões NTFS;

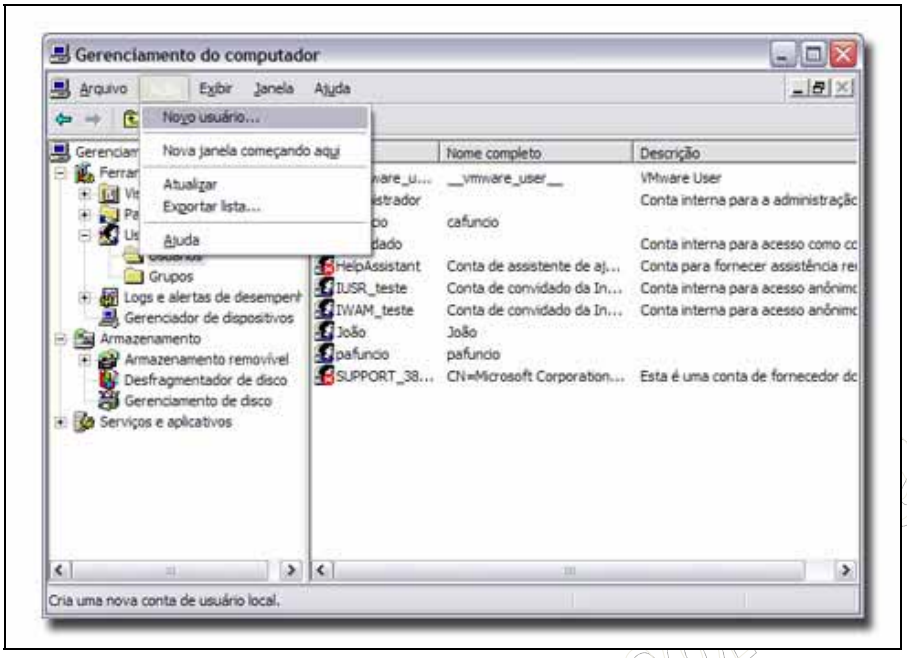
2) Criando grupos e usuários



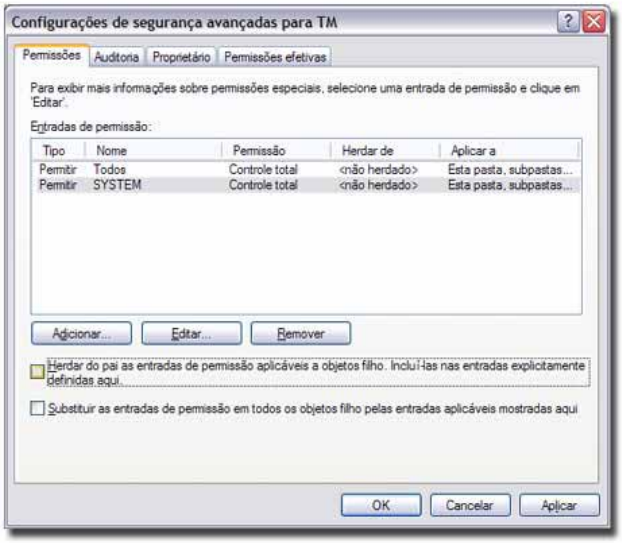


Crie dois grupos : empregados e empregador, e adicione os usuários cafuncio e pafuncio aos empregados, e João ao empregador. Para criar um novo grupo, abra Painel de Controle > Ferramentas Administrativas > Gerenciamento do computador. Clique na árvore de console e, em seguida em Usuários e grupos locais. Clique em Grupos e, em seguida, clique em Ação > Novo grupo

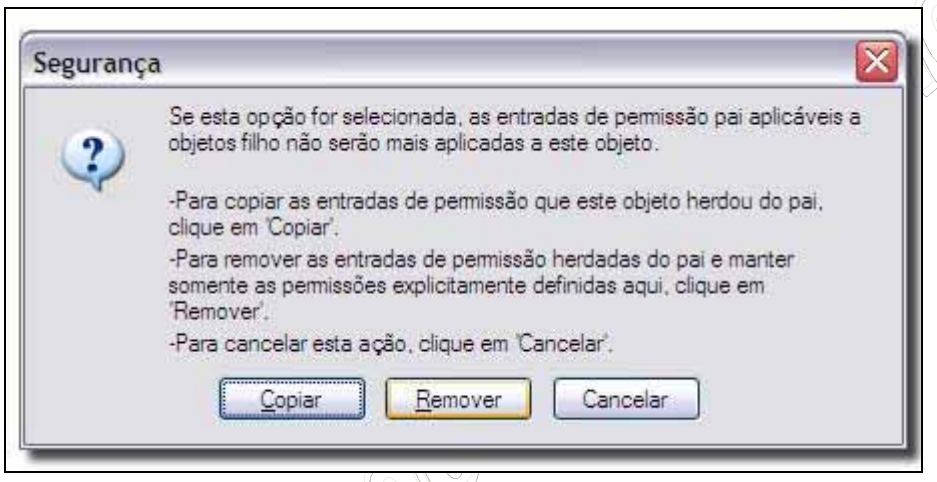
Para criar um novo usuário, siga os passos acima (com uma exceção : na árvore de console, clique em Usuários, e não em Grupos). Adicione cada usuário ao seu respectivo grupo ...



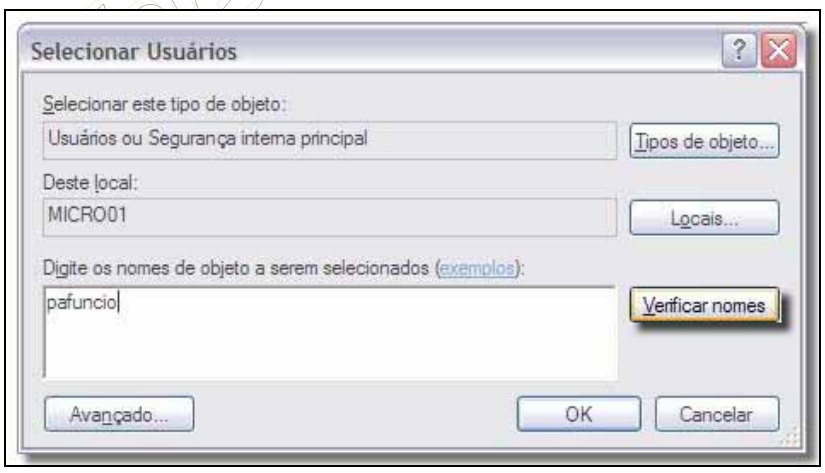
1. Criando permissões: Abra o Windows Explorer e localize a pasta para o qual você deseja definir permissões. Clique com o botão direito do mouse no arquivo ou pasta e abra a folha de Propriedades. Clique na guia Segurança. As permissões NTFS são atribuídas na guia Segurança, e as permissões de compartilhamento na guia Compartilhamento. Se você for membro de um grupo de trabalho e deseja visualizar a guia Segurança, abra o Painel de Controle e clique em Opções de pasta. Na guia Modo de exibição > Configurações Avançadas, desmarque Usar compartilhamento simples de arquivo (recomendável).Observe que todas as caixas de seleção estão sombreadas, e não é permitido modificá-las.
2. O comportamento padrão do Win2000/XP é herdar as permissões do objeto pai (veja Propriedades de objetos > Herança). Há três maneiras de efetuar alterações nas permissões de um objeto:
   - Selecionar a permissões oposta (Negar);
   - Alterar as permissões do objeto pai (as permissões serão herdadas para o objeto filho);
   - Desativar a herança de permissões:



Para desativar as heranças de permissões: Clique em Avançado e na guia Permissões. A coluna Tipo indica o tipo da permissão (Permitir ou Negar); a coluna Nome lista os usuários e os grupos com a respectiva permissão (coluna Permissão). A coluna Herdar de lista a pasta do objeto pai de onde as permissões foram herdadas. A coluna Aplicar a lista as pastas e subpastas às quais uma permissão será aplicada. Desmarque a caixa de seleção Herdar do pai as entradas de permissão aplicáveis a objetos filho. Incluí-las nas entradas explicitamente definidas aqui. Se desejar copiar as permissões herdadas e incluí-las como permissões explícitas (sem herança e que podem ser modificadas pelo administrador), clique em Copiar. O objeto deixará de herdar as permissões do objeto pai. Se deseja removê-las, clique em Remover (somente as permissões explícitas serão mantidas):

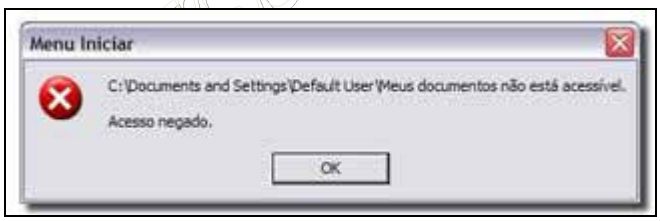


3. Atribua permissões: Clique em OK e retorne. Delete o grupo Usuários (que por padrão é atribuído a qualquer novo usuário criado). Adicione o grupo empregados (atribuir permissões a um grupo é mais rápido e eficiente) ao invés de atribuir aos usuários cafuncio e pafuncio. Atribua as permissões Ler & Executar, Listar Conteúdo da pasta e Leitura:



Adicione o usuário João e atribua permissão de Controle Total : observe que todas as caixas de seleção abaixo são selecionadas:

4. Teste local (no próprio PC): Faça logoff de Administrador (ou de um usuário que tenha permissões de administrador) e faça logon como pafuncio ou cafuncio, que têm as mesmas permissões. Abra o Windows Explorer e localize a pasta para a qual você definiu permissões. Tente acessar a pasta : você conseguirá listar o conteúdo da pasta e subpastas e ler os arquivos (as permissões foram aplicadas a esta pasta, subpastas e arquivos). Tente gravar algum arquivo : Acesso negado ! Os usuários cafuncio e pafuncio pertencem ao grupo empregados, que não tem permissões suficientes para gravar:



Faça logon como João e repita os passos acima. O usuário João tem permissão Controle Total, que permite que grave arquivos na pasta objeto. Esses procedimentos acima podem ser usados em arquivos, pastas e impressoras. As permissões definidas acima são válidas sobre pastas compartilhadas na rede também.

5. Crie agora o compartilhamento da uma pasta ou unidade na rede: Abra o Windows Explorer e localize a pasta que você deseja compartilhar. Clique com o botão direito e em Compartilhamento e segurança... Marque a caixa de seleção **Compartilhar esta pasta na rede** se ela estiver disponível (se ela não estiver disponível, este computador não está em uma rede. Clique no link **Assistente para configuração de rede** e siga as instruções para ativar o compartilhamento de arquivos). Você pode definir o número máximo de usuários acessando o compartilhamento ao mesmo tempo: clique em **Permitir este número de usuários.** As pastas compartilhadas são representadas por uma mão segurando-as:

6. Uma última dica está sobre a criação de compartilhamentos ocultos: Um compartilhamento oculto é quando a pasta está compartilhada, porém você não consegue visualizar através do Explorer. Para criar um compartilhamento oculto: No nome do compartilhamento, digite **$** no último caractere (exemplo : **TM$**, onde **TM** é o nome da pasta). Os recursos compartilhados não podem ser acessados pelo Windows Explorer.

Para acessar um compartilhamento oculto: No menu Iniciar, clique em Executar e digite o comando **UNC** (Convenção universal de nomenclatura). A sintaxe é a seguinte: \\NOME_DO_COMPUTADOR\NOME_DO_COMPARTILHAMENTO\PASTA\NOME_DO_ARQUIVO$, colocando o símbolo “$” ao final do nome doarquivo.

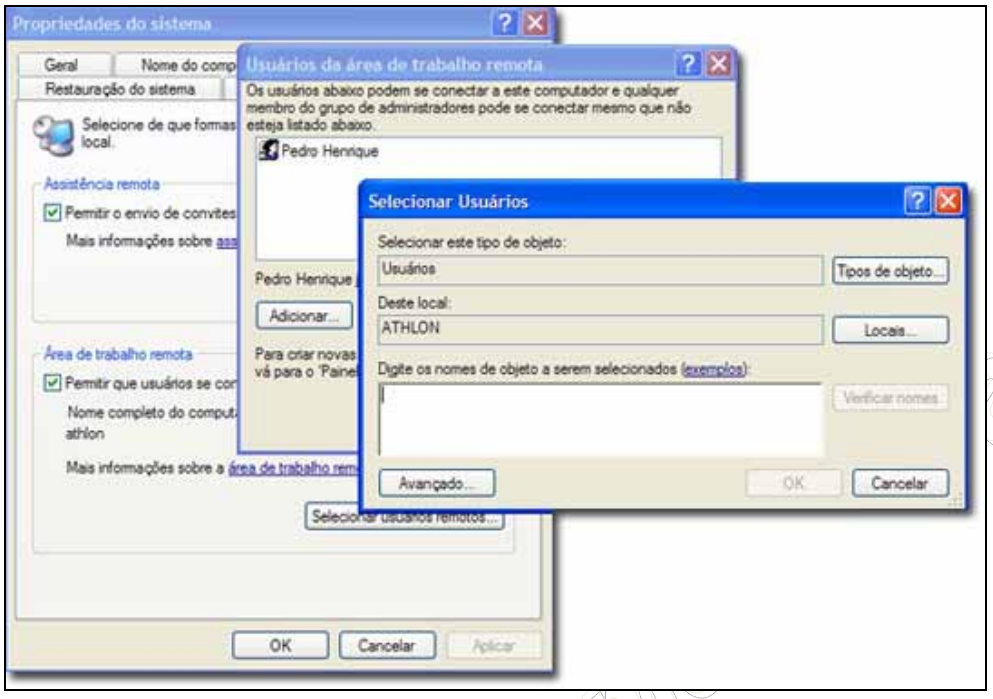
7. Para visualizar todas os compartilhamentos criado em um computador, quem está conectado ao seu computador e quais arquivos estão sendo acessados, execute o seguinte passo: Iniciar -> Painel de Controle -> Ferramentas Administrativas -> Gerenciamento do computador -> Pastas Compartilhadas.

Remote Desktop

O Remote Desktop é um serviço similar ao Terminal Server, porém com recursos limitados. É possível, uma vez habilitado e configurado, acessar uma estação de trabalho remota e ter total controle sobre a mesma. Porém, quando utilizado o Remote Desktop, apenas um único logon é permitido na estação, ou seja, se houver alguém logado na estação, usuários remotos não poderão ter acesso a estação.

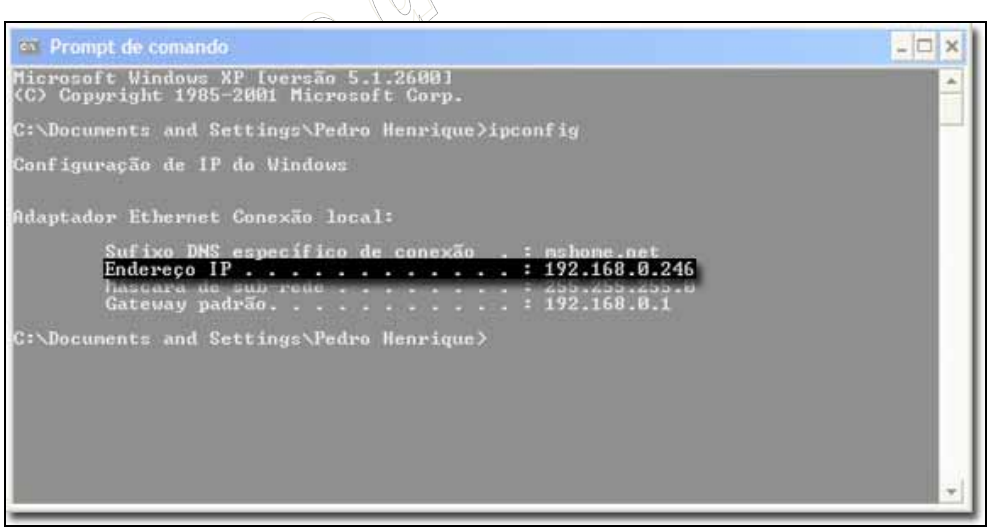
Apesar desta limitação, ainda é considerada uma ferramenta de extrema utilidade. Vamos supor que você tenha saído de sua sala para ir no departamento de RH, mas ao chegar lembra que esqueceu de enviar o e-mail. Você pode acessar remotamento sua estação, que irá efetuar o logoff do seu atual usuário ocioso e então poder enviar o e-mail.

Para habilitarmos o Acesso a Área de Trabalho Remota, realize o seguinte passo: Vá em Iniciar > Painel de controle > Sistema > guia Remoto > marque Permitir que usuários se conectem remotamente à este computador > Selecionar usuários remotos > Adicionar > escreva o nome de algum usuário cadastrado no XP > OK > OK > OK. Pronto.

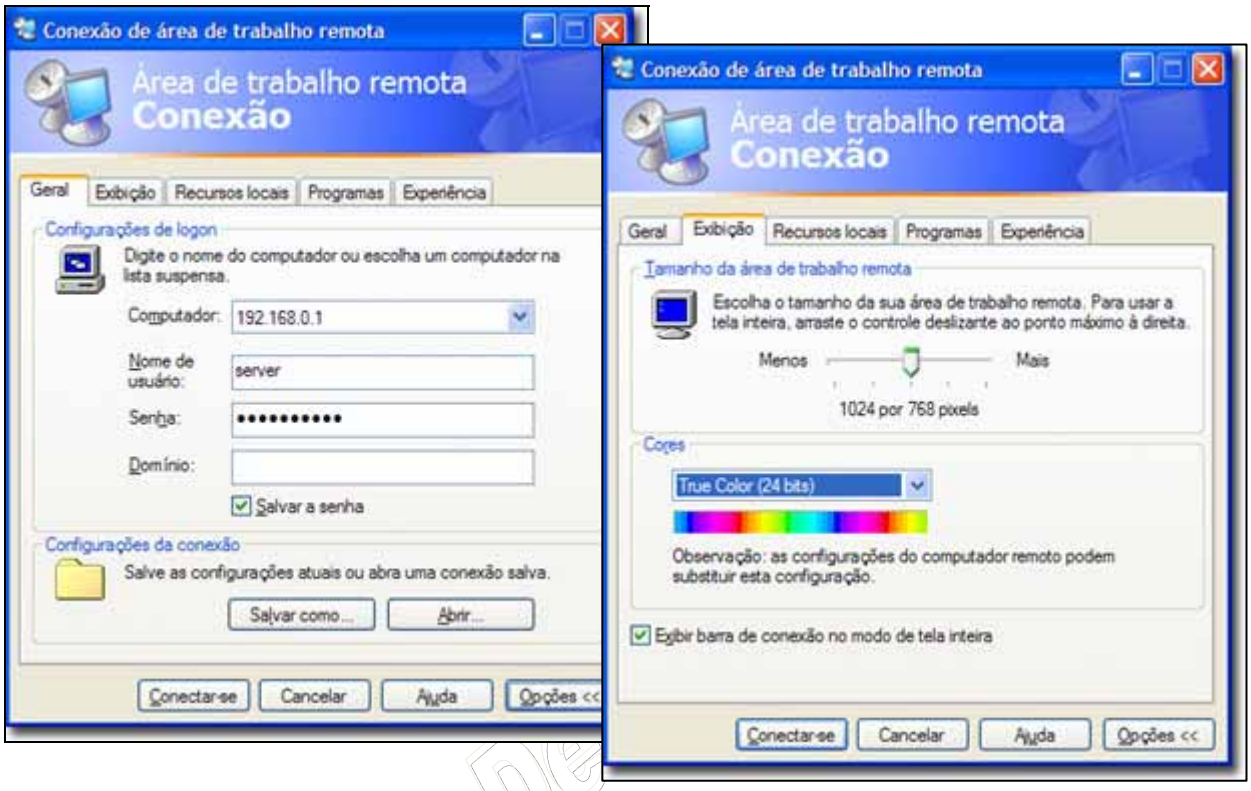


Certifique-se de que a máquina que será acessada não tenha nenhum firewall bloqueando o serviço mstsc.exe (responsável pela Conexão de área de trabalho remota) e de que a máquina esteja conectada diretamente na Internet (sem proxys, servidores, etc).

Ver o IP do computador: Iniciar / Executar / cmd / ipconfig / guarde o número denominado Endereço de IP:

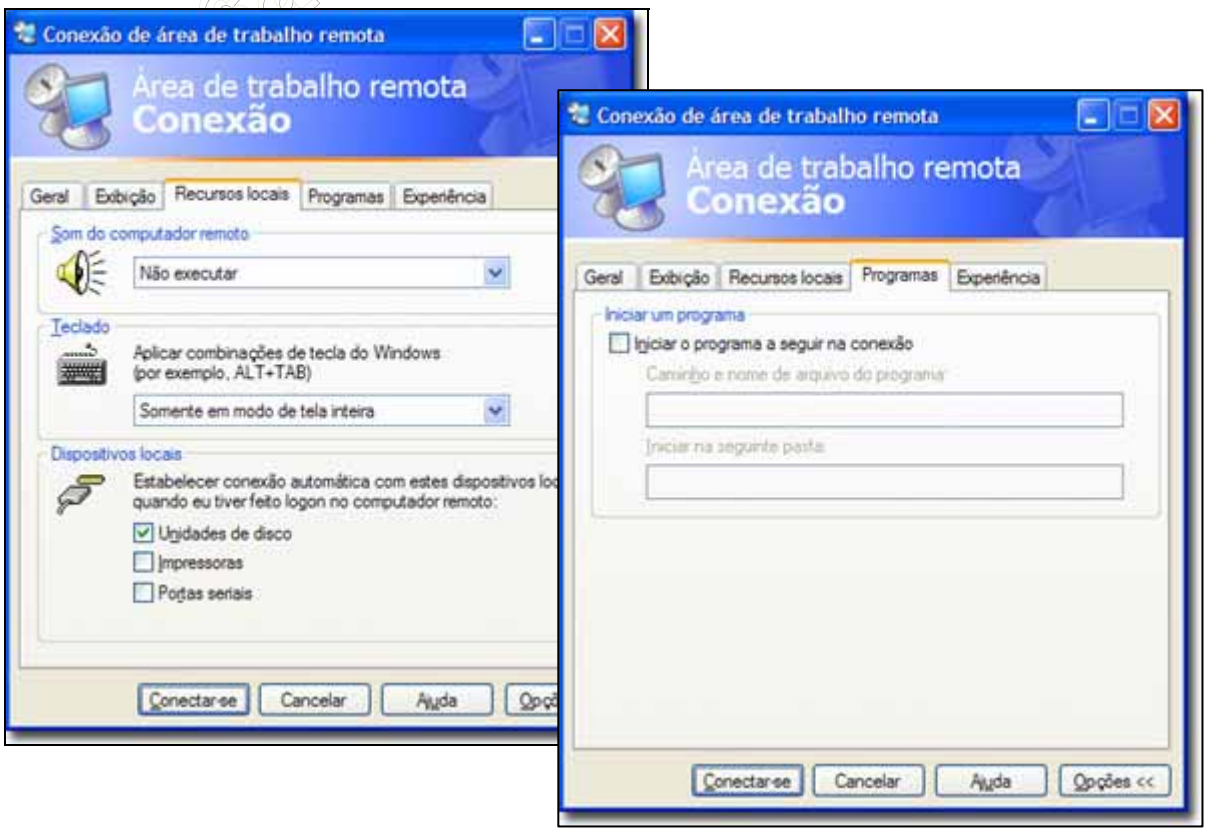


Agora que temos o Remote Desktop habilitado vamos demonstrar a conexão ao mesmo: para isso você pode utilizar o utilitário “mstsc”, o mesmo visto no Terminal Server. Outra alternativa para chegar a este utilitário é: Iniciar > Todos os programas > Acessórios > Comunicações > Conexão de área de trabalho remota. Vejamos algumas das opções desta ferramenta:



Quanto maior a tela e quantidade de cores maior será a necessidade de banda de internet.

Você pode escolher quais recursos locais ficarão visíveis no ambiente virtual e opcionalmente executar um script automático quando efetuar logon:



Por último, você poderá aumentar a velocidade ou desempenho da área remota através da aba "Experiência":

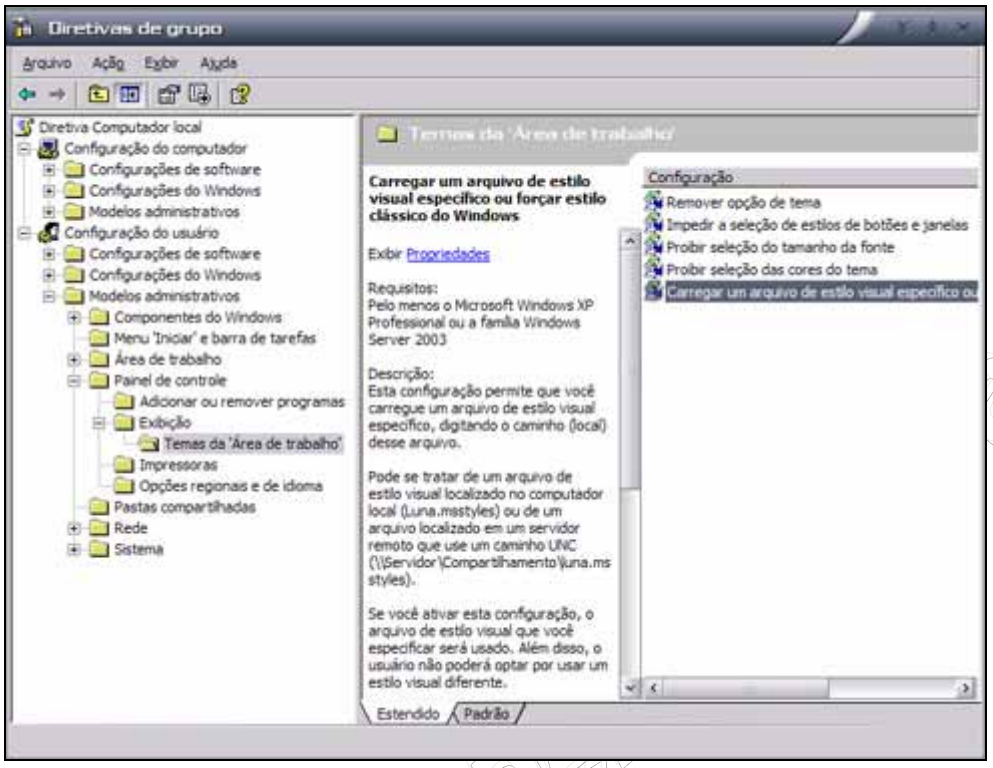
Pronto, você agora já pode realizar a conexão a estação remota. Uma observação é que todas essas configurações feitas não podem ser rejeitadas pela estação remota, como ocorre no Windows Server 2003. Essa é mais uma limitação do Remote Desktop.

Diretivas de Segurança Local

É possível aumentar o nível de segurança e restrições do uso da estação de trabalho através das Diretivas de Grupos ou Diretivas de Segurança Local, no caso de estações de trabalho. É na verdade uma versão mais reduzida e simplificada do que o Group Policie do Active Directory, mas com os mesmo tipos de objetos para edição.

Para acessar as diretivas de segurança local: Iniciar -> Painel de Controle -> Ferramentas Administrativas -> Diretivas de Segurança Local. Ou como alternativa: Iniciar -> Executar -> gpedit.msc

Aqui você poderá realizar as mesmas restrições vista no GPM, mas para os usuários locais da estação de trabalho:



Após realizar suas configurações é necessário reiniciar o computador.

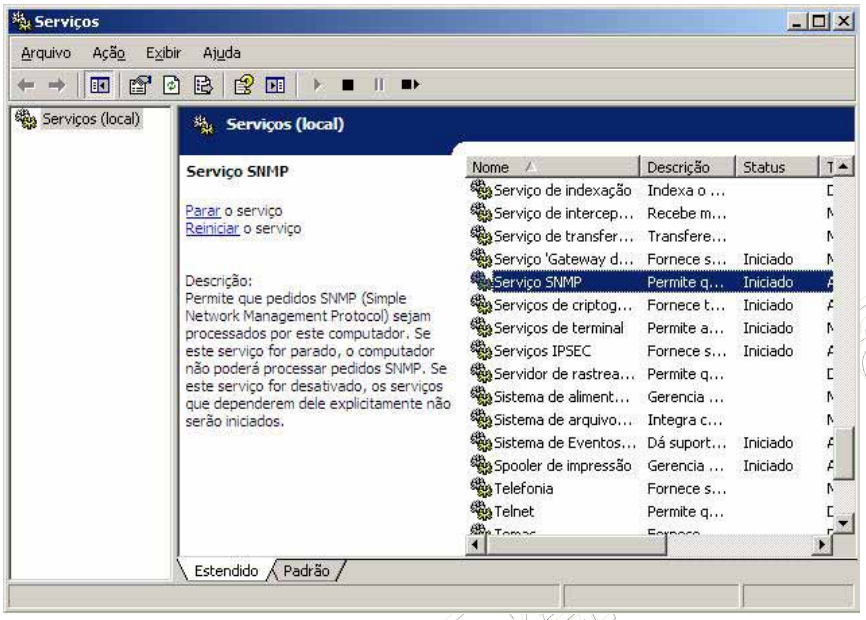
11.5 GERENCIAMENTO DE REDES

A gerência de redes envolve basicamente: o equipamento que se deseja gerenciar e a estação de gerência de rede. Vejamos primeiramente como preparar uma estação de trabalho Windows para ser gerenciada através da rede, e em seguida como utilizar um software de gerenciamento da rede.

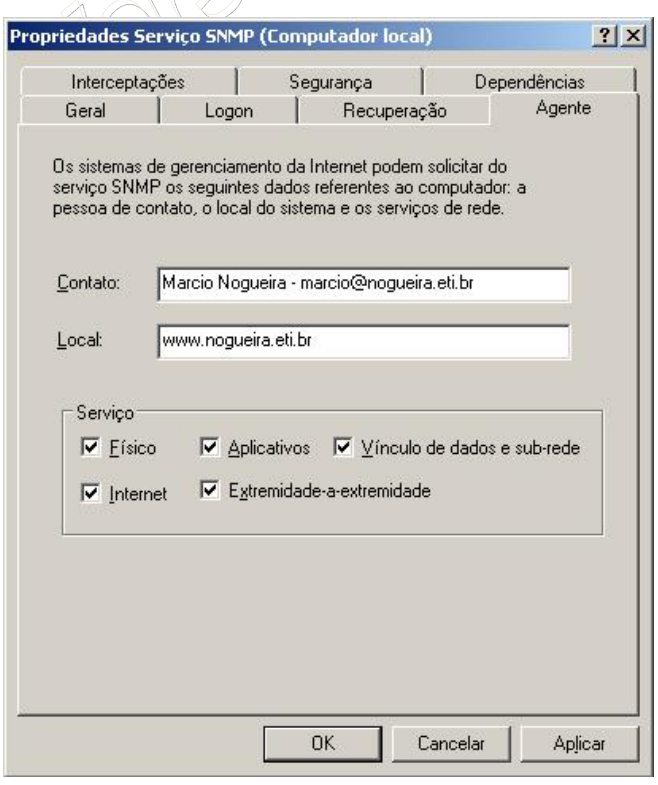
Para instalar o Serviço SNMP no Windows você precisará:

1. Clique no botão **Iniciar** e, em seguida, vá para o **Painel de Control**e e execute **Adicionar ou remover programas**. Na caixa de diálogo clique em **Adicionar / Remover Componentes do Windows**;
2. Nos componentes do Windows, clique sobre **Ferramentas de Gerenciamento**, e clique em **Detalhes**;
3. Marque a caixa de **Simple Network Management Protocol (SNMP)**;
4. É provável que você seja solicitado pelo CD de instalação do Windows, dessa forma mantenha-o próximo para o caso disso ocorrer;

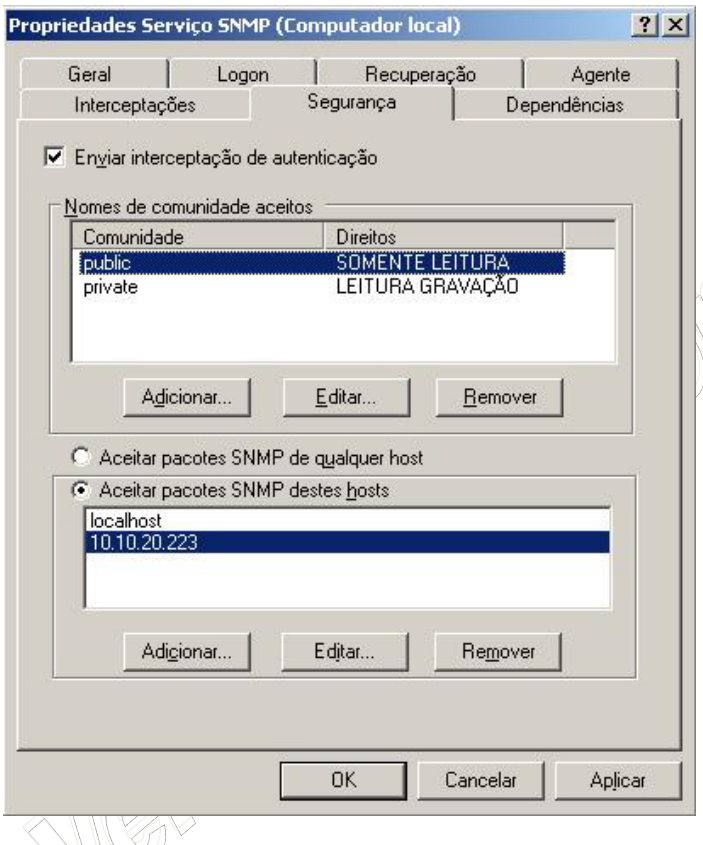
5. O Serviço SNMP será instalado e iniciado automaticamente no Windows, porém é preciso configurar seus parâmetros e segurança, para isso volte ao **Painel de Controle**, depois vá em **Ferramentas Administrativas**, e clique em **Serviços**;



6. Dois novos serviços foram adicionados ao Windows, o SNMP Service, que é o principal motor de monitoramento e acesso pela estação de gerência da rede, e o Trap SNMP, que é responsável por coletar atividades no seu computador e enviá-las para a estação de gerenciamento. Clique sobre o **SNMP Service** e acesse a aba de **propriedades**;
7. Nas propriedades, clique em **Agente** é configure suas informações de **contato** e **localidade**, e em seguida selecione os **serviços** que deseja monitorar com o SNMP.



8. Por último, clique na aba de **Segurança**. Aqui você deverá configurar as chamadas **comunidades**, que equivale a configuração de autenticação da estação monitora da rede com sua estação de trabalho. Por padrão, o SNMP reconhece os seguintes tipos de comunidades: public (que permite um acesso por parte da estação monitora de rede de apenas leitura), e private (que permite o aceito e modificação de informações no computador por parte da estação monitora da rede).



Recomendamos que você escolha comunidades diferentes das padrões, do contrário, qualquer outra pessoa que descubra essas comunidades poderão acessar suas informações ou modificar suas configurações. Porém não esqueça de anotá-las, pois precisaremos delas para configurar a estação monitora da rede mais adiante.

Em seguida, configure os hosts (estações de monitoramento) que terão permissão para acessar a sua estação de trabalho.

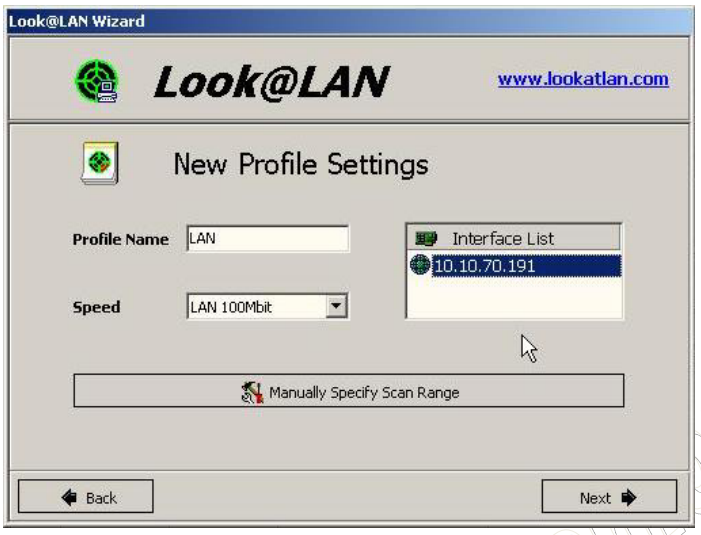
Por padrão ele trás a configuração de somente permitir que sua própria estação de trabalho possa acessar, para verificar o funcionamento do serviço. Configure um segundo endereço IP, o endereço da sua estação de monitoramento da rede, neste mesmo campo.

Agora que temos o agente SNMP instalado, configurado e em execução na estação de trabalho, vamos aprender como utilizar um software de gerenciamento de redes.

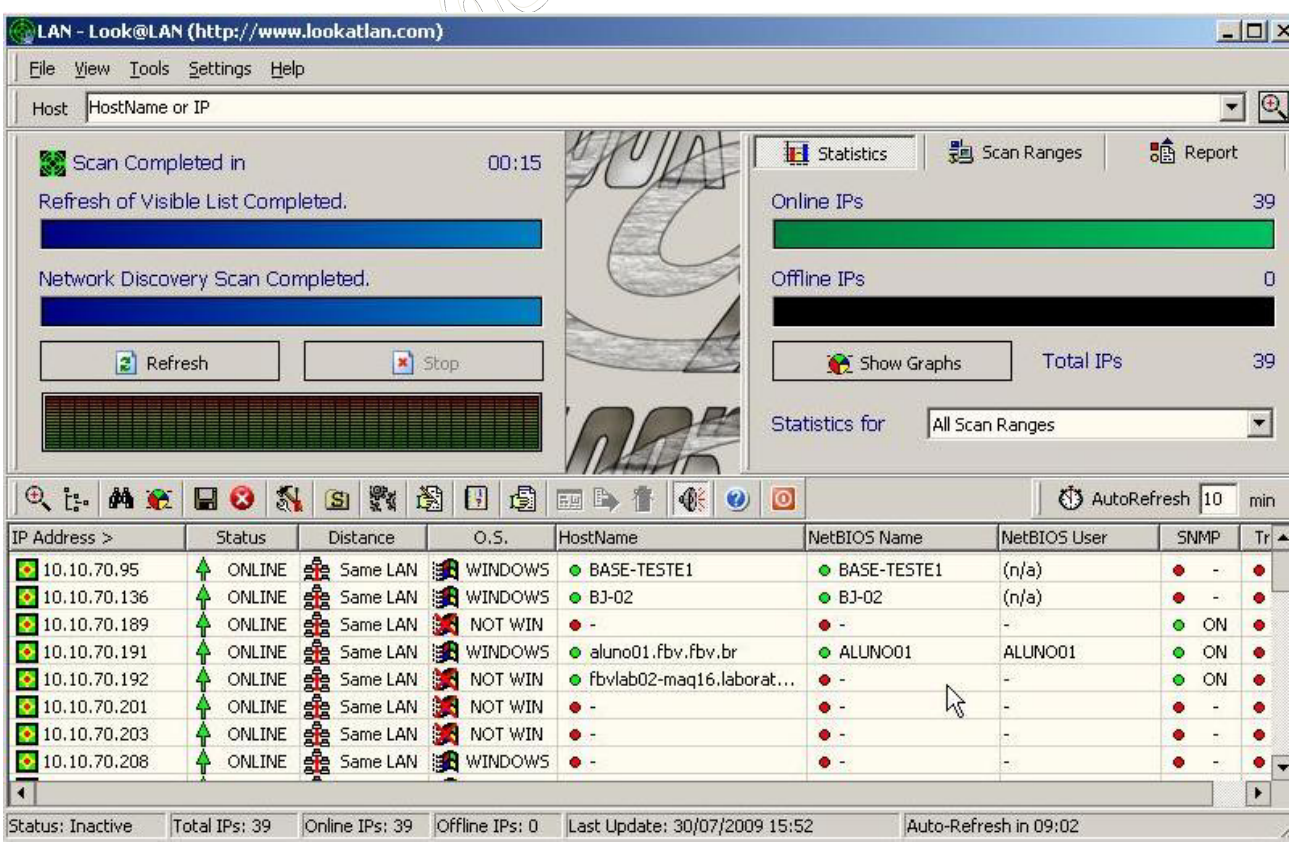
Optamos em utilizar o software livre Look@Lan Network Monitor, desenvolvido pelo estudante Carlo Medas, como forma de exemplificar a usabilidade do SNMP em rede, diversos outros softwares estão a disposição, inclusive algumas versões comerciais oferecem suporte ao envio de mensagens via e-mails, sms, alerta sonoro, desenvolvimento de planta baixa, entre outros.

Ao executarmos o Look@Lan pela primeira vez será solicitado para criarmos um perfil, este perfil é uma forma de memorizar os dados que serão retornados pelos Agentes SNMP

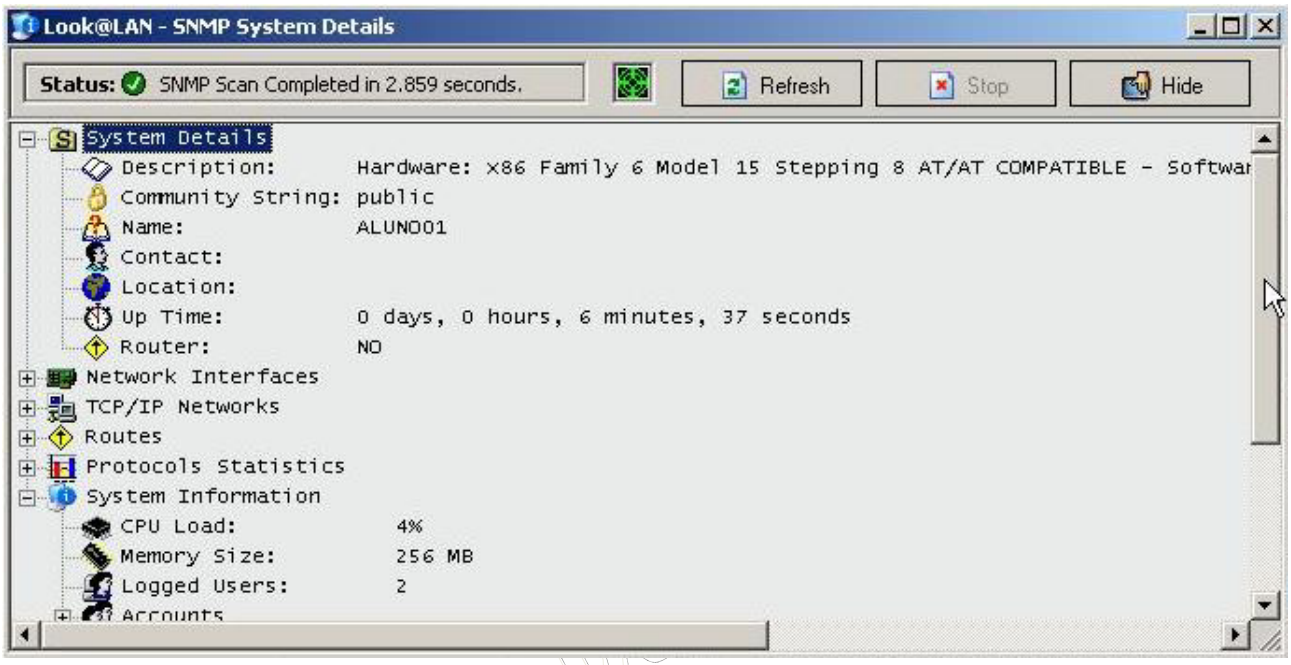
Vamos optar em analisar toda a rede a fim de identificar quais estações de trabalho estão com o Agente SNMP habilitado:



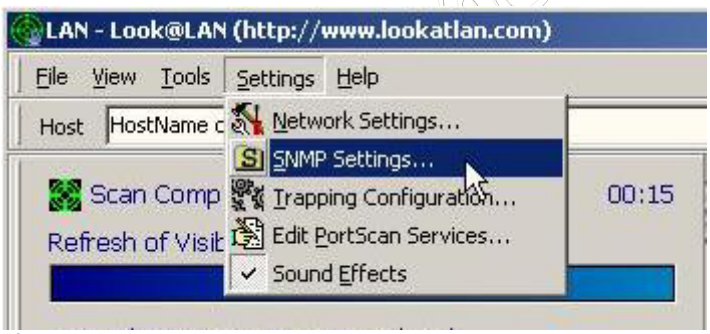
Ao término da varredura, o software apresentará uma relação de hosts que estão ativos na rede, e na coluna (SNMP) os que estiverem com o status em ON implica nas estações cujo agente SNMP está em atividade:



Quando clicamos sobre uma estação, cujo SNMP esteja ON, será apresentado diversas informações sobre a estação. As informações exibidas podem variar de estação para estação, conforme a configuração do agente SNMP.

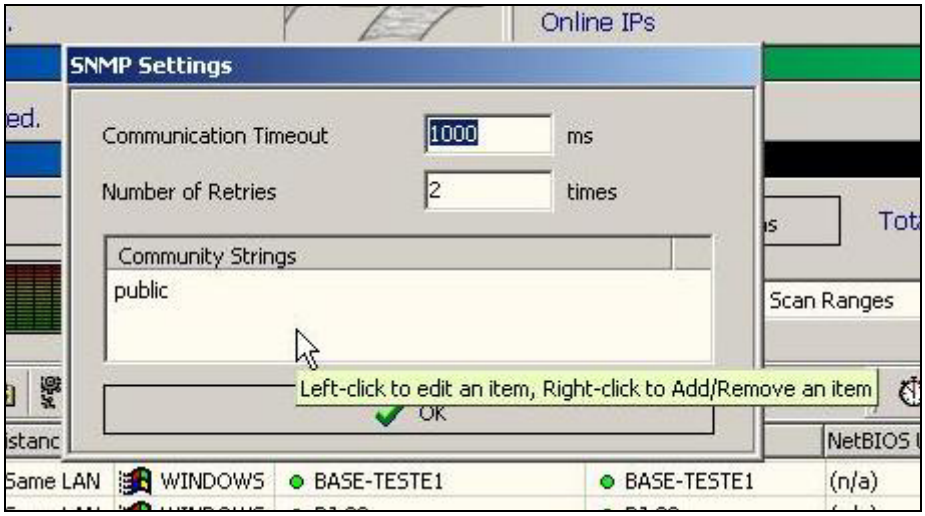


Para finalizar, vejamos a configuração do SNMP neste software de gerenciamento de Rede.



Acesse a aba de **Settings**, e em seguida **SNMP Settings**.

Será apresentada uma caixa para inclusão da Community (comunidade). Aqui você deverá listar todas as comunidades que estão configuradas em suas estações, do contrário o software não conseguirá acesso via o agente SNMP.



12 CONSIDERAÇÕES FINAIS

A presente apostila é resultado da preparação para o curso *In Company* a ser realizado na CHESF, em Agosto de 2009. Seu conteúdo é fruto de um compêndio realizado sobre livros de redes, em especial aos livros de:

1. SOARES, Luiz Fernando Gomes; LEMOS, Guido; COLCHER, Sergio; Redes de computadores - das LANs, MANs e WANs as redes ATM.. Rio de Janeiro: Campus, 1999.
2. TANENBAUM, Andrew S; Redes de computadores. Traducao Vanderberg Dantas de Souza . Rio de Janeiro: Campus, 2003.

Monografias diversas, localizadas através do Google e do Google Acadêmico, sites públicos na Internet, em especial:

1. Clube do Hardware – www.clubedohardware.com.br
2. Wikipedia – pt.wikipedia.org
3. Infowest – www.infowest.com.br

E contribuições pessoais do professor Márcio L. M. Nogueira, ao longo de toda a obra.

Os direitos pessoais e de imagem não foram citados por se tratar de uma apostila acadêmica, não publicada, não distribuída e sem fins comerciais de editoração e venda da obra.

Qualquer tipo de uso, reprodução parcial, reprodução total ou citação fora do curso específico é considerado uma prática ilegal. Bem como sua utilização em outros cursos ou turmas do mesmo curso. Percebendo o uso, solicitamos a gentileza de comunicar ao Prof. Márcio Nogueira, através do e-mail de contato: marcio@nogueira.eti.br, a fim de que as devidas medidas sejam tomadas.

Este material não substitui as devidas literaturas e deve ser utilizado única e exclusivamente com fins de suporte a sala de aula.