

Atualmente, a quantidade de dados que as empresas precisam guardar dobra a cada 12 a 18 meses. As tecnologias de armazenamento secundário são os discos magnéticos, os discos ópticos, as fitas magnéticas e as redes de armazenamento.

Disco magnético O meio de armazenamento secundário mais amplamente utilizado hoje é o **disco magnético**. Os PCs possuem **discos rígidos**, e os grandes mainframes ou sistemas de computador de médio porte contam com múltiplos drives de discos rígidos, pois requerem imensa capacidade de armazenagem em disco, na faixa dos gigabytes e terabytes. Alguns PCs mais antigos também usam discos flexíveis, mas hoje em dia essa tecnologia foi largamente suplantada pelos **drives removíveis USB**, também conhecidos como drives USB. Um drive USB oferece memória de armazenamento instantânea e portátil, bastando conectá-lo a uma porta USB do computador. Eles proporcionam mais de seis gigabytes de capacidade de armazenamento portátil e são tão pequenos que cabem no bolso.

Computadores grandes, com imensas necessidades de armazenamento, usam uma tecnologia de disco denominada **Raid** — *Redundant Array of Inexpensive Discs*. Os dispositivos Raid formam pacotes que reúnem mais de uma centena de drives de discos, um chip controlador e um software especializado em uma única unidade de grande porte, que apresenta os dados por múltiplas vias simultaneamente.

Discos ópticos Os discos ópticos utilizam a tecnologia de laser para armazenar quantidades maciças de dados, incluindo som e imagem, de forma altamente compacta. Existem discos ópticos para PCs e para computadores de grande porte. O sistema de disco óptico mais comum utilizado com PCs é chamado **CD-ROM** (*compact disk read-only memory* — **disco compacto com memória somente de leitura**). Um disco compacto de 4,75 polegadas para PCs pode guardar até 660 megabytes. O CD-ROM é um armazenamento somente para leitura, mas os discos **CD-RW** (*CD-ReWritable* — **CD regravável**) são regraváveis. Os **discos de vídeo digital (DVDs)** são discos ópticos do mesmo tamanho dos CD-ROMs, mas com capacidade mínima para 4,7 gigabytes de dados. Os DVDs são usados para armazenar vídeos e dados digitalizados de textos, recursos gráficos e áudio. Os discos regraváveis (DVD-RW) são amplamente utilizados nos sistemas computacionais pessoais.

Fita magnética Algumas empresas ainda utilizam **fita magnética**, tecnologia de armazenamento ainda mais antiga, empregada para armazenamento secundário de grandes quantidades de dados que precisam ser acessados com rapidez, mas não instantaneamente. Ela armazena dados sequencialmente e é lenta, comparada à velocidade de outros meios de armazenamento secundário.

Redes de armazenamento de dados Para lidar com suas necessidades crescentes de armazenamento, as grandes organizações estão recorrendo a tecnologias de armazenamento em rede. As **redes de armazenamento de dados** (*storage area networks* — **SANs**) conectam inúmeros dispositivos de armazenamento em uma rede de alta velocidade independente e dedicada. A SAN cria um grande grupo central de armazenamento rapidamente acessado e compartilhado por múltiplos servidores (veja a Figura 4.4).

Dispositivos de entrada e saída

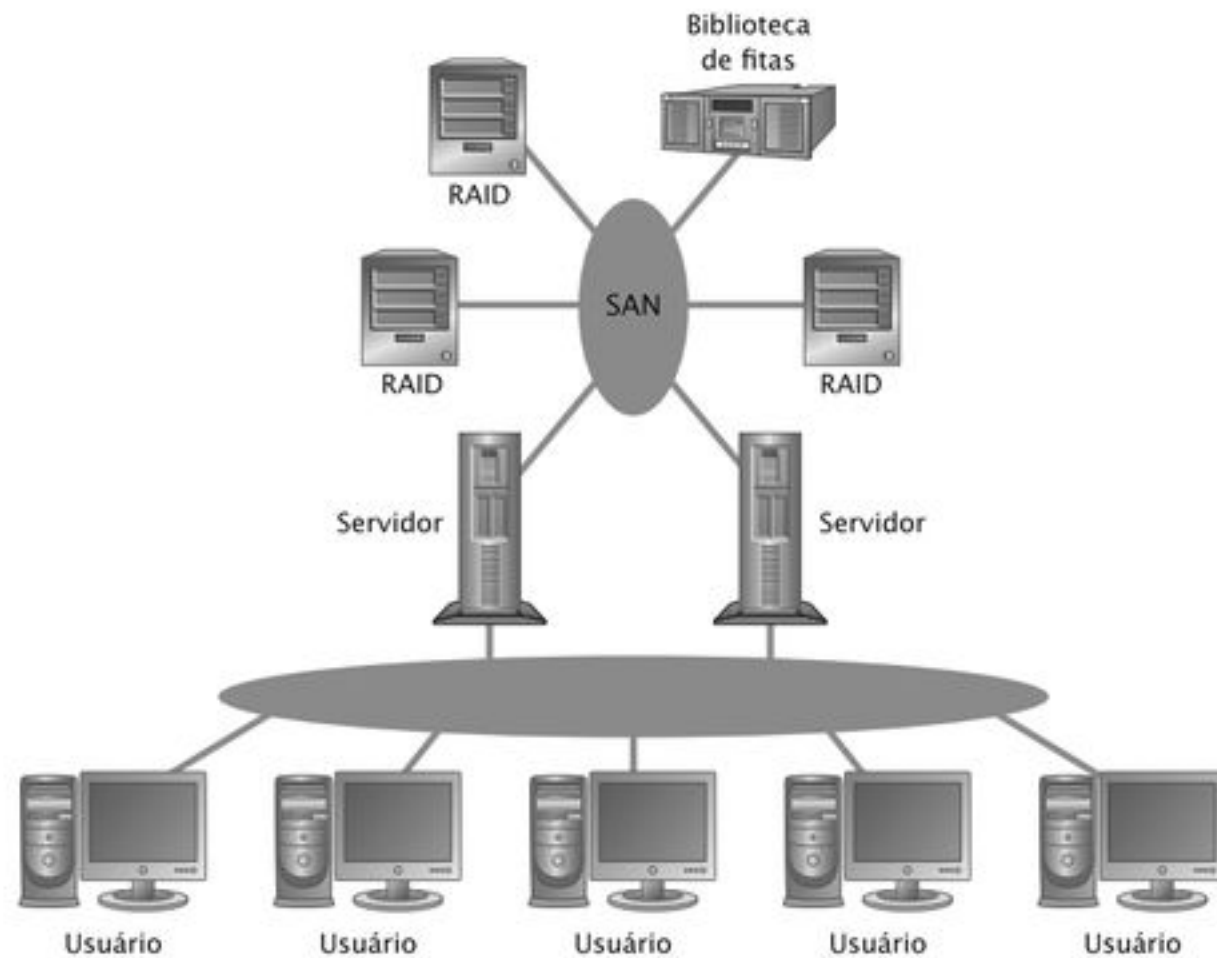
A interação humana com sistemas de computador acontece geralmente por meio de dispositivos de entrada e saída. **Dispositivos de entrada** recolhem dados e os convertem em formato eletrônico para uso pelo computador, ao passo que **dispositivos de saída** apresentam os dados depois de processados. Na Tabela 4.1 são descritos os principais dispositivos de entrada e saída.

Tendências contemporâneas de hardware

A crescente capacidade do hardware e da tecnologia de rede vem mudando radicalmente a maneira como as empresas organizam sua capacidade computacional; cada vez mais, elas confiam partes desse volume às redes. Atualmente, observamos seis tendências: plataforma

Figura 4.4**Rede de armazenamento de dados (SAN)**

A SAN armazena dados em diversos tipos diferentes de dispositivos, provendo dados para a empresa. Ela apoia a comunicação entre qualquer servidor e a unidade de armazenamento, bem como entre os diferentes dispositivos de armazenamento da rede.



digital móvel emergente, nanotecnologia, computação em nuvem, computação autônoma, virtualização e processadores multinúcleo.

Plataforma digital móvel emergente

O Capítulo 1 ressaltou que novas plataformas computacionais digitais móveis surgiram como alternativa aos PCs e grandes computadores. Dispositivos de comunicação, como celulares e smartphones, a exemplo do iPhone e do BlackBerry, assumiram muitas funções dos PDAs, incluindo transmissão de dados, utilização da Internet, transmissão de e-mails e mensagens instantâneas, exibição de conteúdo digital e troca de dados com sistemas corporativos internos. A nova plataforma móvel também inclui notebooks pequenos, leves e de baixo custo chamados **netbooks**, otimizados para comunicação sem fio e acesso à Internet, com funções computacionais importantes como o processamento de texto, além de novos leitores digitais com alguns recursos de acesso à Web, como o Kindle, da Amazon, e os tablets, como o iPad.

Um volume cada vez maior de computação empresarial está sendo transferido de PCs e máquinas desktop para esses dispositivos móveis. Por exemplo, os executivos seniores da General Motors norte-americana utilizam aplicações de smartphones que fornecem informações sobre venda de veículos, desempenho financeiro, métricas de produção e status da gestão de projetos. Na AstraTech, fabricante de aparelhos médicos, os representantes de venda utilizam seus smartphones para acessar aplicações e dados de vendas do sistema Salesforce.com de gestão do relacionamento com o cliente, consultando informações sobre produtos vendidos e devolvidos e tendências gerais de receita antes de se encontrarem com os clientes. Os empregados da Kraft Foods nos Estados Unidos usam iPhones para e-mail e lista de contatos, e para acesso a documentos e wikis relacionados a projetos através do servidor Microsoft SharePoint da empresa.

Nanotecnologia

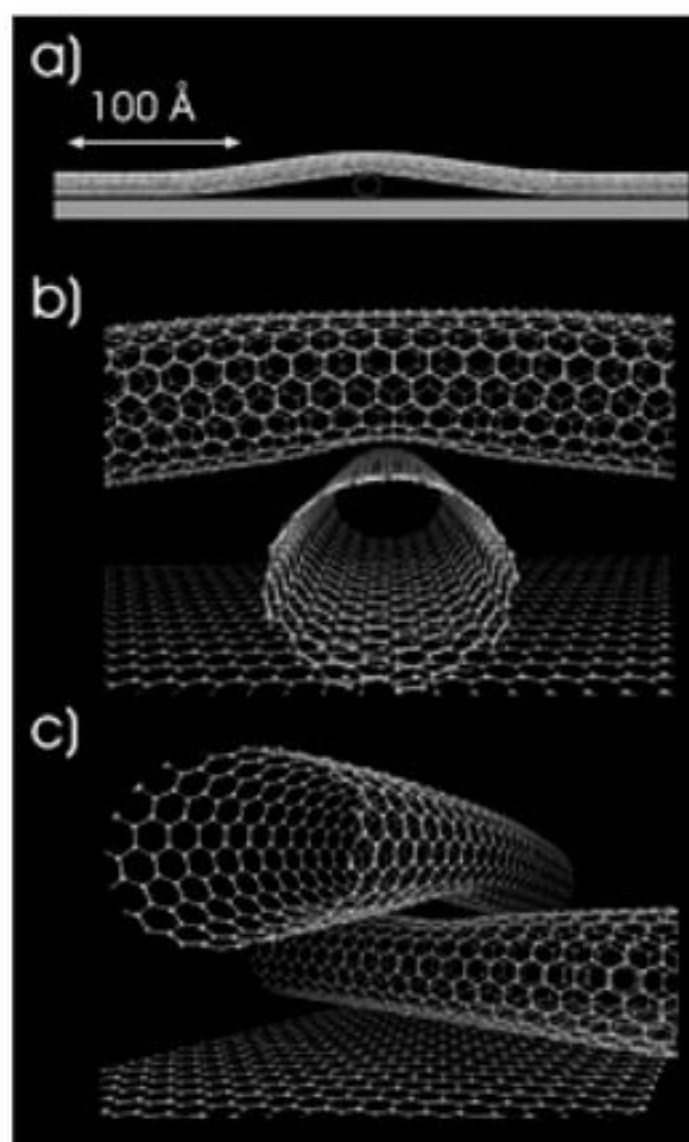
Ao longo dos anos, os fabricantes de microprocessadores conseguiram aumentar exponencialmente o poder de processamento enquanto diminuíam o tamanho do processador, encontrando maneiras de incluir mais transistores em um espaço cada vez menor. Agora, eles recorrem à **nanotecnologia** para diminuir o tamanho dos transistores para o de diversos átomos. A nanotecnologia utiliza átomos e moléculas individuais para criar processadores de computador e outros dispositivos milhares de vezes menores do que o tamanho da

Tabela 4.1 Dispositivos de entrada e saída

<u>Dispositivo de entrada</u>	<u>Descrição</u>
Teclado	Principal método de entrada para dados numéricos e de texto.
Mouse de computador	Dispositivo manual com capacidade de apontar e clicar geralmente ligado ao computador por um cabo. O usuário pode movimentar o mouse sobre uma mesa para controlar a posição do cursor na tela do monitor e apertar um botão para selecionar um comando. No caso de PCs do tipo laptop, muitas vezes são usados dispositivos de toque (<i>touch pads</i>) e <i>track balls</i> no lugar do mouse tradicional.
Tela sensível ao toque	Dispositivo que permite aos usuários inserir quantidades limitadas de dados tocando a superfície sensibilizada de um monitor de vídeo com o dedo. Encontrada frequentemente em quiosques de informação instalados em aeroportos, lojas de varejo, restaurantes e em smartphones multitoque como o iPhone e o iPad.
Reconhecimento óptico de caracteres	Dispositivo que converte marcas, caracteres e códigos especialmente projetados para formato digital. O código óptico mais usado é o código de barras utilizado em sistemas de pontos de venda de supermercados e lojas de varejo. Os códigos podem ter dados de horário, data e localização, além dos de identificação.
Reconhecimento de caracteres de tinta magnética (MICR – <i>magnetic ink character recognition</i>)	Tecnologia utilizada principalmente no setor bancário para processamento de cheques. Os caracteres na parte inferior de um cheque identificam o banco, a conta e o número do cheque, e são previamente impressos com uma tinta magnética especial. A MICR converte esses caracteres ao formato digital para processamento por computador.
Entrada por caneta	Dispositivos de reconhecimento de escrita à mão, como pranchetas, agendas e blocos de anotação com canetas, que convertem para formato digital os movimentos feitos por uma caneta eletrônica pressionada sobre uma tela sensível ao toque.
Dispositivo de varredura digital (scanner digital)	Dispositivo que converte imagens, como figuras e documentos, para formato digital. É um componente essencial dos sistemas de processamento de imagens.
Entrada de áudio	Dispositivos de entrada por voz que convertem palavras faladas em formato digital para processamento por computador. Microfones e players podem servir como dispositivos de entrada para música e outros sons.
Sensores	Dispositivos que coletam dados diretamente do ambiente para entrada em um sistema de computador. Por exemplo, os fazendeiros hoje podem usar sensores para monitorar a umidade do solo em suas lavouras e tomar decisões sobre irrigação.
<u>Dispositivo de saída</u>	<u>Descrição</u>
Monitores	Tela composta por um painel plano ou, em sistemas antigos, tubo de raios catódicos (CRT – <i>cathode Ray tube</i>).
Impressoras	Dispositivos que produzem cópia impressa da informação de saída. Compreendem impressoras de impacto (como as de matriz de pontos) e impressoras sem impacto (laser, jato de tinta e transferência térmica).
Saídas de áudio	Dispositivos de saída de voz que reconvertem dados digitais de saída para fala inteligível. Outras saídas de áudio, como música, podem ser reproduzidas por caixas de som conectadas ao computador.

Figura 4.5**Exemplos de nanotubos**

Nanotubos são tubos pequeninos, cerca de 10 mil vezes mais finos do que um fio de cabelo humano. Eles consistem de folhas de carbono em forma de hexágono enroladas; podem potencialmente ser utilizados como fios minúsculos ou em dispositivos eletrônicos ultrapequenos e em poderosos condutores de corrente elétrica.



tecnologia atual. A IBM e outros laboratórios de pesquisa criaram transistores a partir de nanotubos e outros dispositivos elétricos e desenvolveram um processo de manufatura para produção econômica de processadores nanotubo (Figura 4.5).

Computação em nuvem

O termo 'computação em nuvem' refere-se a um modelo de computação no qual empresas e indivíduos obtêm recursos computacionais e aplicações de software pela Internet (também conhecida como 'a nuvem'). Milhares, ou mesmo centenas de milhares, de computadores estão localizados em centros de dados na nuvem, onde podem ser acessados por computadores desktop, notebooks, netbooks, centros de entretenimento, dispositivos móveis e outras máquinas clientes conectadas à Internet. IBM, HP, Sun Microsystems, Dell e Amazon operam centros de computação em nuvem gigantescos e escaláveis para oferecer tanto poder computacional quanto armazenamento de dados e conexões de alta velocidade à Internet para empresas que desejam manter sua infraestrutura de TI remotamente. Empresas de software como Google, Microsoft, SAP, Oracle e Salesforce.com comercializam aplicações, como serviços oferecidos pela Internet. A Figura 4.6 ilustra o conceito de computação em nuvem.

A computação em nuvem é formada por três tipos diferentes de serviços:

- **Infraestrutura em nuvem como serviço** Para rodar seus sistemas de informação, os clientes utilizam processamento, armazenamento, redes e outros recursos computacionais fornecidos por provedores de serviço de nuvem. A Amazon, por exemplo, utiliza a capacidade excedente de sua infraestrutura de TI para oferecer uma infraestrutura de TI para serviços de venda largamente baseada no ambiente em nuvem. Isso inclui o serviço *Simple Storage Service (S3)*, para armazenamento de dados de clientes, e o serviço *Elastic Computing Cloud (EC2)*, para execução de sua aplicação. Os usuários pagam somente pelo montante da capacidade de processamento e armazenamento que realmente utilizam.

**Figura 4.6****Plataforma de computação em nuvem**

Na computação em nuvem, as capacidades de hardware e software são oferecidas como serviços na Internet. Utilizando um dispositivo conectado à Internet, empresas e empregados têm acesso a aplicações e infraestrutura de TI em qualquer lugar e a qualquer hora.

- **Plataforma em nuvem como serviço** Os clientes utilizam a infraestrutura e as ferramentas de programação hospedadas pelo provedor de serviços para desenvolver suas próprias aplicações. A Sun Microsystems, por exemplo, oferece os serviços *Sun Storage Cloud* e *Sun Compute Cloud* para ajudar os desenvolvedores de software, estudantes e iniciantes a testar e desenvolver novas aplicações baseadas na Internet utilizando hardware da Sun. A IBM possui um serviço semelhante, o *Smart Business Application Development & Test*, para desenvolvimento e teste de software de nuvem.
- **Software em nuvem como serviço** Os clientes utilizam software hospedado pelo fornecedor. Exemplos pioneiros são o Google Apps, que oferece aplicativos comuns on-line, e o Salesforce.com, que disponibiliza CRM e outros serviços de software relacionados pela Internet. Ambos cobram dos usuários uma taxa anual de assinatura, embora o Google Apps ofereça uma versão equivalente gratuita. Os usuários acessam essas aplicações a partir de um navegador da Web, e dados e software são mantidos nos servidores remotos dos provedores. Pode-se saber mais sobre a Salesforce.com e seus serviços no estudo de caso ao final do capítulo, e sobre computação em nuvem no Capítulo 11.

Como as organizações que utilizam computação em nuvem normalmente não possuem a infraestrutura, elas não precisam fazer grandes investimentos em hardware e software próprios. Em vez disso, elas adquirem seus serviços computacionais junto a provedores remotos e pagam somente pelo montante de poder computacional que utilizam (computação de utilidade) ou são cobradas mensal ou anualmente, conforme plano de assinatura. O termo **computação sob demanda** também tem sido usado para descrever tais serviços.

A computação em nuvem possui algumas desvantagens. A menos que os usuários tenham planos de armazenar seus dados localmente, a responsabilidade pelo armazenamento e controle deles fica nas mãos do provedor. Algumas empresas se preocupam com os riscos de segurança relacionados à confiança de seus dados e sistemas críticos a um fornecedor externo que também trabalha com outras empresas. Existem ainda questões relacionadas à confiabilidade do sistema. As empresas esperam que seus sistemas estejam disponíveis em tempo integral e não querem sofrer qualquer perda na capacidade de negócios se sua infraestrutura de TI funciona mal. Em julho de 2008, quando a nuvem da Amazon se desconectou, os usuários não conseguiram usar seus sistemas por oito horas. Outra limitação da computação em nuvem é a dependência dos usuários aos serviços do provedor.

Existem pessoas que acreditam que a computação em nuvem representa uma transformação na maneira como a computação será realizada pelas empresas à medida que a computação

empresarial migra de centros de dados privados para 'a nuvem' (Carr, 2008). Essa questão continua a ser tema de debates. A princípio, a computação em nuvem é mais atrativa para as pequenas e médias empresas que não dispõem dos recursos para adquirir e manter hardware e software próprios. Entretanto, grandes empresas possuem fortes investimentos em sistemas proprietários complexos que dão apoio a processos de negócio exclusivos, com alguns conferindo vantagens estratégicas à empresa. Para eles, o cenário mais provável é um modelo de computação híbrido, nos quais as empresas utilizam sua própria infraestrutura para as atividades principais e adotam a computação em nuvem para os sistemas menos críticos ou para a capacidade de processamento adicional durante os períodos de pico nos negócios. A computação em nuvem fará com que as empresas gradualmente migrem da capacidade de infraestrutura fixa para uma infraestrutura mais flexível, com algumas partes pertencendo à empresa e outras alugadas dos gigantescos centros computacionais pertencentes aos fornecedores de hardware.

Computação autônoma

Com grandes sistemas compostos por milhares de dispositivos em rede, os sistemas computacionais se tornaram tão complexos que alguns especialistas acreditam que, no futuro, eles serão simplesmente inadministráveis. Uma abordagem para lidar com esse problema sob a perspectiva do hardware é utilizar a computação autônoma. A **computação autônoma** é uma iniciativa setorial para desenvolver sistemas capazes de configurar, otimizar e sintonizar a si mesmos, autoconsertar-se quando avariados e proteger-se de intrusos e da autodestruição. Imagine, por exemplo, um PC de mesa que pudesse perceber quando é invadido por um vírus. Em vez de permitir-se inadvertidamente ser invadido, o PC identificaria e erradicaria a ameaça ou, como alternativa, repassaria sua carga de trabalho a outro processador e se desligaria automaticamente antes que o vírus destruísse qualquer arquivo.

Você pode ter uma ideia de alguns desses recursos no seu computador pessoal. Por exemplo, o software de firewall e antivírus pode detectar vírus nos PCs, eliminá-los automaticamente e alertar os operadores. Esses programas são atualizados de modo automático conforme a necessidade conectando-se a um serviço antivírus on-line, como o McAfee. A computação autônoma ainda carece de outros elementos-chave, que certamente estarão na pauta dos centros de pesquisa nesta década. Ela acontece quase que diariamente quando a Microsoft, a Apple e a Sun atualizam automaticamente os computadores de seus usuários quando eles estão conectados à Internet.

Virtualização e processadores multinúcleo

À medida que as empresas instalam centenas ou milhares de servidores, elas gastam com eletricidade para alimentar e refrigerar seus sistemas quase o mesmo que gastaram com os equipamentos. A Agência de Proteção Ambiental norte-americana (EPA – *Environmental Protection Agency*) estimou que, até 2011, os centros de dados irão utilizar mais de dois por cento de toda a energia elétrica dos Estados Unidos. Acredita-se que a tecnologia de informação contribuirá com dois por cento dos gases mundiais do efeito estufa. A diminuição do consumo de energia nos centros de dados tornou-se não só um problema sério, como também um desafio ambiental. A Seção Interativa sobre organizações examina esse problema. Ao ler o caso, tente identificar as soluções alternativas para o problema e as vantagens e desvantagens de cada uma delas.

Essa Seção Interativa descreve organizações que contiveram a proliferação de hardware e o consumo de energia utilizando a virtualização para diminuir o número de computadores necessários no processamento. A **virtualização** apresenta uma série de recursos computacionais (como poder computacional e armazenamento de dados) de modo que todos possam ser acessados de maneira que não estejam restritos pela configuração física ou pela localização geográfica. A virtualização de servidores permite que as empresas rodem mais de um sistema operacional ao mesmo tempo em uma única máquina. A maioria dos servidores utiliza somente de 10 a 15 por cento de sua capacidade, e a virtualização pode aumentar a taxa de utilização dos servidores para 70 por cento ou mais. Taxas de utilização mais altas representam uma menor necessidade de computadores para processar o mesmo montante de trabalho.

SEÇÃO INTERATIVA: ORGANIZAÇÕES**A computação verde é boa para os negócios?**

As salas de computadores estão ficando muito quentes. Tarefas de alto consumo de dados — como vídeo sob demanda, cópia de músicas da Internet, troca de fotos e manutenção de sites — requerem um número cada vez maior de computadores que consomem muita energia. Entre os anos de 2000 e 2007, o custo total anual de eletricidade para os servidores de centros de dados subiu de 1,3 bilhão de dólares para 2,7 bilhões de dólares nos Estados Unidos, e de 3,2 bilhões de dólares para 7,2 bilhões de dólares ao redor do mundo. Se a tendência persistir, a eletricidade total utilizada por servidores em 2010 deve ser 76 por cento maior do que em 2005, conforme estudo de Jonathan Kooney, cientista do Lawrence Berkeley National Laboratory. Os consultores da Gartner Group acreditam que as contas de energia elétrica, responsáveis por 10 por cento do orçamento para tecnologia de informação, podem chegar a mais de 50 por cento em breve.

O calor gerado pelas salas cheias de servidores faz com que os equipamentos parem de funcionar. As empresas são forçadas a gastar ainda mais com a refrigeração de seus centros de dados ou a encontrar outras soluções. Algumas organizações gastam mais dinheiro com a refrigeração de seus centros de dados do que com o aluguel da propriedade. É um ciclo vicioso, já que as empresas precisam pagar para alimentar seus servidores e, então, pagar novamente para mantê-los resfriados e operacionais. A refrigeração de um servidor demanda quase o mesmo número de quilowatts de energia que seu funcionamento. Todo esse consumo adicional de energia tem impacto negativo sobre o meio ambiente e os custos operacionais da empresa.

Algumas das empresas mais importantes do mundo enfrentam questões relacionadas ao consumo de energia atentas ao meio ambiente e à diminuição de gastos. Google, Microsoft e HSBC estão construindo centros de dados que irão se beneficiar de energia hidrelétrica. A Hewlett-Packard está trabalhando em uma série de tecnologias para reduzir em 75 por cento a pegada de carbono pelos centros de dados e para desenvolver novos serviços e software para medir o uso de energia e as emissões de carbono. A empresa reduziu entre 20 e 25 por cento seus custos de energia através da consolidação de servidores e centros de dados.

O centro de dados da Microsoft em San Antonio dispõe de sensores que medem quase todo o consumo de energia, recicla a água utilizada na refrigeração e usa um software para gestão da energia desenvolvido internamente. A organização também tenta encorajar práticas de economia de energia por meio da cobrança às unidades de negócios com base na quantidade de energia que consomem no centro de dados, em vez de cobrar pelo espaço que ocupam no andar.

Nenhuma dessas empresas diz que seus esforços vão salvar o mundo, mas elas demonstram reconhecer um problema crescente e o começo da era da computação verde. E, como as tecnologias e os processos dessas empresas são mais eficientes do que os de outras

empresas, utilizar seus serviços de software on-line em vez do software *in-house* também é considerado um investimento verde.

Diariamente, os PCs costumam ficar ligados pelo dobro do tempo que realmente são usados. Segundo um relatório da Alliance to Save Energy, uma empresa com 10 mil PCs desktop gastará mais do que 165 mil dólares por ano com contas de energia elétrica se as máquinas permanecerem ligadas durante a noite. O grupo estima que, somente nos Estados Unidos, essa prática desperdiça cerca de 1,7 bilhão a cada ano.

Embora muitas empresas estabeleçam configurações para a gestão de energia de PCs, cerca de 70 por cento dos empregados desativam essas configurações. O 1E NightWatchman, software de gestão de energia da BigFix, e o Verdiem travam as configurações de energia do PC e ligam automaticamente os computadores pela manhã pouco antes de os empregados chegarem para o trabalho.

As escolas públicas de Miami-Dade County diminuíram o tempo em que os PCs permaneciam ligados de 21 horas para 10,3 horas por dia utilizando a BigFix para controlar centralmente as configurações de energia dos computadores. A City University of New York adotou o software Surveyor, da Verdiem, para desligar seus 20 mil PCs quando estão inativos durante a noite. O software diminuiu em 10 por cento os gastos com energia da universidade, criando uma economia anual de cerca de 320 mil dólares.

A virtualização é uma ferramenta altamente eficaz para uma computação verde mais econômica, pois reduz o número de servidores e recursos de armazenamento na infraestrutura de TI da empresa. Fulton County, na Geórgia, oferece serviços para 998 mil cidadãos e avalia o uso de energia quando adquire novas tecnologias de informação. A região utilizou o software de virtualização VMWare e a nova plataforma de servidor blade Fujitsu para consolidar servidores legados subutilizados de modo que, agora, uma máquina executa o trabalho que costumava ser realizado por oito máquinas, economizando 44 mil dólares por ano nos gastos com energia. Esses esforços também criaram uma infraestrutura de TI mais atualizada.

A virtualização também encoraja a consolidação de pessoas e processos. Segundo estudo da International Data Center realizado em 2006, gastos gerenciais e administrativos crescem três vezes mais rápido do que os gastos com equipamentos computacionais. A virtualização facilita o lançamento de novas aplicações em servidores existentes e reduz problemas associados a determinado servidor físico. Há menos servidores para supervisionar, embora ainda precisem ser cuidadosamente gerenciados e monitorados.

Especialistas observam que é importante que as organizações meçam seu estoque e uso de energia e controlem seus ativos de tecnologia de informação tanto antes quanto depois das iniciativas verdes. Métricas comumente usadas e empregadas pela Microsoft e

outras empresas incluem uso eficiente de energia, eficiência da infraestrutura dos centros de dados e eficiência média dos centros de dados. A seguradora de saúde Highmark inicialmente desejava aumentar a utilização de sua CPU em 10 por cento e reduzir o consumo de energia de 5 a 10 por cento. Quando a empresa inventariou todos seus ativos de tecnologia de informação, descobriu que a equipe de sistemas de informação estava conservando servidores 'mortos', que não faziam nada além de consumir energia.

Programas para educar empregados com relação à conservação de energia também podem ser necessários. Além de utilizar ferramentas para a monitoração, a

Honda Motor Corporation treina seus administradores de centros de dados a serem mais eficientes em termos de utilização de energia. A empresa ensinou-os, por exemplo, a desativar rapidamente os equipamentos não utilizados e a utilizar ferramentas de gestão para garantir que os servidores sejam otimizados.

Fontes: Jim Carlton, "The PC goes on an energy diet". *The Wall Street Journal*, 8 set. 2009; Ronan Kavanagh, "IT virtualization helps to go green". *Information Management Magazine*, mar. 2009; Ross O. Storey, "Don't forget the green IT imperative". *MIS Asia*, 23 mar. 2009; J. Nicholas Hoover, "10 ideas to power up your green IT agenda". *Information Week*, 22 set. 2008; e Eric Chabrow, "The wild, wild cost of data centers". *CIO Insight*, maio 2008.

PERGUNTAS SOBRE O ESTUDO DE CASO

1. Que problemas sociais e empresariais são causados pelo consumo de energia nos centros de dados?
2. Quais as soluções disponíveis para esses problemas? Quais delas estão preocupadas com o meio ambiente?
3. Quais são os benefícios empresariais e os custos dessas soluções?
4. Todas as empresas deveriam seguir na direção da computação verde? Justifique.

A rede de hospitais e cuidados da saúde Christus Health, por exemplo, presente nas regiões sul e oeste dos Estados Unidos e no México, costumava gerenciar mais de 2 mil servidores em oito centros de dados, com 70 por cento deles localizado em San Antonio. Lá, 97 por cento dos sistemas utilizavam 20 por cento, ou menos, de sua capacidade de processamento e somente 29 por cento da memória disponível. A empresa de assistência médica usou a virtualização para consolidar o trabalho de 824 servidores em 83 servidores *blade*, economizando 1,8 milhão de dólares que incluíam a redução do consumo de energia.

O software de virtualização do servidor funciona entre o sistema operacional e o hardware, escondendo dos usuários os recursos do servidor, inclusive o número e a identidade de servidores físicos, processadores, e sistemas operacionais. O VMWare é o fornecedor de software de virtualização líder para os sistemas Windows e Linux. A Microsoft incluiu recursos de virtualização na última versão do Windows Server.

Além de reduzir os gastos com hardware e energia, a virtualização permite que as empresas executem suas aplicações legadas em versões mais antigas de um sistema operacional instalado no mesmo servidor das novas aplicações. A virtualização também facilita a centralização da administração de hardware.

Processadores multinúcleo Outra maneira de reduzir os requisitos de energia e o amontoamento de hardware é utilizar processadores multinúcleo. Um **processador multinúcleo** é um circuito integrado ao qual dois ou mais processadores foram conectados para a melhoria do desempenho, redução do consumo de energia e processamento simultâneo mais eficiente de múltiplas tarefas. Essa tecnologia permite que duas máquinas de processamento com requisitos reduzidos de consumo de energia e dissipação de calor executem tarefas de modo mais rápido do que um núcleo com somente um processador. Atualmente, é possível encontrar processadores *dual-core* (dois núcleos) em PCs e *quad-core* (quatro núcleos) em servidores. O processador UltraSparc T2, da Sun Microsystems, gerencia cargas de trabalho na Internet e possui oito processadores. A Intel está trabalhando em um processador com 80 núcleos.

Infraestrutura de TI: software

Para usar o hardware, você precisará do software. É ele que oferece instruções detalhadas para guiar o trabalho do computador. O conceito de software inclui software de sistema e software aplicativo, já definidos neste capítulo. Esses dois tipos de software estão inter-rela-

cionados e podem ser pensados como um conjunto de caixas alojadas uma dentro da outra, cada uma delas interagindo com as demais. A Figura 4.7 ilustra essa relação. O software de sistema engloba e controla o acesso ao hardware. Para operar, o software aplicativo precisa trabalhar por meio do software de sistema. Os usuários finais trabalham primordialmente com softwares aplicativos. Projetando cada tipo de software para uma máquina específica, garante-se a compatibilidade.

Software de sistema operacional

O software de sistema que gerencia e controla as atividades do computador é denominado **sistema operacional**. Outro software de sistema é composto de programas tradutores de linguagens computacionais que convertem as linguagens de programação em linguagem de máquina compreensível pelo computador, e de programas utilitários que executam tarefas de processamento comuns, tais como copiar, classificar ou calcular uma raiz quadrada.

O sistema operacional é o gerente-geral do sistema de computador, permitindo que esse sistema lide com várias tarefas e usuários ao mesmo tempo. Ele aloca e designa recursos do sistema, programa a utilização dos recursos e tarefas e monitora as atividades do sistema. Além disso, provê locais na memória primária para dados e programas e controla os dispositivos de entrada e saída, como impressoras, terminais e conexões de comunicação. O sistema operacional também coordena a programação das tarefas em execução no computador, de modo que partes diferentes de tarefas diversas possam ser executadas ao mesmo tempo. Finalmente, acompanha cada tarefa do computador e ainda monitora quem está usando o sistema, os programas executados e quaisquer tentativas não autorizadas de acesso ao sistema.

Sistemas operacionais de PCs, de servidores e móveis

O sistema operacional controla a forma como um usuário interage com o computador. Os sistemas operacionais e muitos tipos de aplicações atuais usam **interface gráfica de usuário (graphical user interface)**, muitas vezes chamada **GUI**, que faz uso extensivo de ícones, botões, barras e caixas para executar tarefas.

Novas tecnologias de interface estão surgindo tanto para os sistemas empresariais quanto para os domésticos. Uma tecnologia de interface promissora é a **multitoque**, popularizada pelo iPhone /iPad. A Seção Interativa sobre tecnologia explora as interfaces multitoque como alternativa às interfaces GUI. Ao ler o caso, tente identificar os problemas a serem resolvidos pelas interfaces e as questões relacionadas a pessoas, organizações e tecnologia a serem tratadas pela solução.

A Tabela 4.2 compara os principais sistemas operacionais para PCs e servidores. Ela inclui a família de sistemas operacionais Windows (Windows 7, Windows Vista, Windows Server 2008), além do UNIX, do Linux e do sistema operacional do Macintosh.

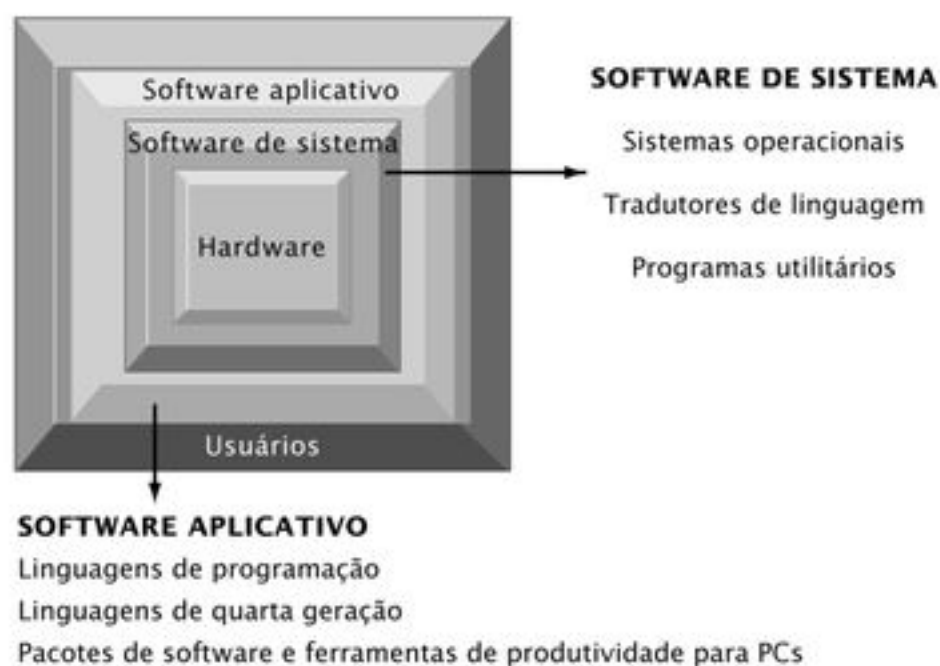


Figura 4.7

Os principais tipos de software

A relação entre software de sistema, software aplicativo e usuários pode ser ilustrada por uma série de caixas alojadas uma dentro da outra. O software de sistema — composto de sistemas operacionais, tradutores de linguagem e programas utilitários — controla o acesso ao hardware. Para operar, os softwares aplicativos, como as linguagens de programação e as de ‘quarta geração’, devem trabalhar por meio do software de sistema. O usuário interage primariamente com o software aplicativo.

SEÇÃO INTERATIVA: TECNOLOGIA A novidade do toque

Quando Steve Jobs demonstrou pela primeira vez 'o beliscão' — o movimento com dois dedos para aumentar e diminuir o zoom de fotos e páginas da Web no iPhone —, ele não impressionou somente a indústria de telefonia móvel, mas também todo o mundo digital. Os recursos multitoque do iPhone da Apple demonstraram novas maneiras de utilizar o toque para interagir com programas e dispositivos.

As interfaces de toque não são novas. As pessoas fazem uso delas diariamente para sacar dinheiro nos caixas eletrônicos ou para realizar o *check-in* em quiosques de aeroportos. Pesquisadores comerciais e acadêmicos trabalham em tecnologias multitoque há anos. O que a Apple fez foi tornar o recurso mais atraente e relevante, popularizando-o assim como fez com o mouse e a interface GUI na década de 1980. (Esses recursos também foram inventados em outro lugar.)

As interfaces multitoque são potencialmente mais versáteis do que a interface de um único toque. Elas permitem que um ou mais dedos sejam utilizados para executar gestos especiais que manipulam listas ou objetos em uma tela sem que se mova o mouse, clique em botões, role telas ou pressione teclas. Elas executam ações distintas dependendo da quantidade de dedos que detectam e do movimento que o usuário realiza.

A tela multitoque do iPhone e seu software permitem que se controle tudo utilizando somente os dedos. Um painel embaixo do vidro da tela sente o seu toque utilizando campos elétricos. Em seguida, transmite a informação para a tela de LCD abaixo dele. Um software especial reconhece múltiplos pontos simultâneos de toque, em oposição às telas de toque único, que reconhecem somente um ponto de toque. Pode-se alternar rapidamente entre uma série de páginas da Web e fotos tocando a tela com maior intensidade ou posicionando três dedos sobre a tela e deslizando-os para qualquer um dos lados. Beliscando a imagem, é possível diminuí-la ou aumentá-la.

Outras empresas estão levando produtos multitoque ao mercado. A Synaptics, fornecedor líder de *touchpads* para fabricantes de laptop que concorre com a Apple, anunciou que está incorporando diversos recursos multitoque em seus *touchpads*. A Microsoft lançou recentemente seu computador *Surface*, que roda Windows 7 e permite que seus clientes empresariais usem multitoque em uma tela sobre uma mesa. Clientes de hotéis, cassinos e lojas de varejo poderão utilizar gestos multitoque para mover-se entre objetos digitais tais como fotos, divertir-se com jogos e navegar entre opções de produtos. É possível que a tecnologia do *Surface* seja integrada aos PCs. O *tablet PC* Latitude XT, da Dell, utiliza multitoque; o que é útil para aqueles que não conseguem manipular um mouse ou desejam a funcionalidade de um PC tradicional. No lugar do mouse, é possível usar os dedos ou uma caneta apontadora.

A tela multitoque do BlackBerry Storm usa algo denominado 'toque háptico'. Quando pressiona um botão na tela ou no teclado virtual, a sensação é a de

que você está realmente tocando aquele ponto específico. É palpável, o que lhe deixa mais tranquilo se sofre de problemas auditivos. Se tem problemas visuais, isso torna muito mais fáceis a digitação e os cliques.

A Hewlett-Packard (HP) agora possui laptops e desktops que utilizam a tecnologia de toque. Seu computador *TouchSmart* permite que sejam utilizados dois dedos de uma vez para manipular imagens na tela ou para realizar movimentos sobre a tela que indicam comandos específicos sem que se recorra a cursores ou barras de rolagem. Quando é possível colocar o dedo diretamente sobre a tela, não é necessário um cursor mostrando o local onde está apontando. Para mover um objeto, você toca sobre ele com um dedo e o arrasta para sua nova localização. Deslizar o dedo para cima e para baixo ou para os lados suavemente faz a tela rolar.

O *TouchSmart* permite que os usuários domésticos embarquem em um novo tipo de computação casual — deixando mensagens escritas, de vídeo ou de áudio aos familiares, consultando rapidamente direções antes de sair de casa, ouvindo música enquanto prepara o jantar. Consumidores e empresas também já encontraram outros usos. Segundo Alan Reed, vice-presidente da HP e gerente-geral da Business Desktops, "existe um potencial ainda não descoberto de tecnologia de toque no mercado empresarial para inserir os usuários em uma forma nunca antes experimentada".

O aeroporto O'Hare, de Chicago, integrou um conjunto de PCs *TouchSmart* nos quiosques 'Explore Chicago', permitindo que os visitantes percorressem um centro de visitantes virtual. A tecnologia *TouchSmart* ajudou um estudante autista a falar e se comunicar com os outros pela primeira vez em 14 anos de vida. Sem utilizar o teclado e o mouse sem fio do PC *TouchSmart*, os usuários podem realizar bate-papos por vídeo com empregados remotos através de uma webcam e um microfone integrados, acessar e-mail e Internet, e gerenciar contatos, itens de calendário e fotos.

Os PCs que permitem o toque também poderiam ser úteis em escolas de ensino fundamental em busca de um computador de fácil manuseio para os alunos das séries iniciais, ou buscando dispositivo de informações do tipo quiosque, montado na parede, para pais e visitantes. Os sistemas de toque também permitem que clientes se conectem e interajam com fornecedores e outros clientes. É possível ainda que os clientes usem o toque para fazer pedidos a um varejista, realizar chamadas virtuais de serviço por vídeo, operar ou ensinar a utilizar redes sociais nos negócios.

O novo sistema operacional Windows 7, da Microsoft, ostenta recursos multitoque: quando conecta uma tela de toque a esse sistema, é possível consultar jornais on-line, folhear álbuns de fotografias e organizar pastas e arquivos somente com os dedos. Para aproximar o zoom de algo, basta colocar os dois dedos sobre a tela de um PC multitoque compatível e os afastar. Para clicar com o botão direito sobre um arquivo, toque-o com um dedo e bata na tela com outro.

É muito cedo para saber se a nova tela multitoque algum dia será tão grandiosa quanto a interface gráfica orientada a mouse. Embora colocar os dedos sobre a tela seja a medida mais atual para o que é 'legal' em termos de celulares, uma 'excelente aplicação de toque' para os PCs ainda não foi inventada. Já é evidente, entretanto, que o toque possui vantagens reais em dispositivos nos quais não é possível ou conveniente usar um mouse, ou em interfaces muito antigas de menus e pastas já obsoletas.

Fontes: Kathy Sandler, "The future of touch". *The Wall Street Journal*, 2 jun. 2009; Ashlee Vance, "PC touch screens move ahead". *The New York Times*; Suzanne Robitaille, "Multitouch to the rescue?". *Suite101.com*, acesso em 22 jan. 2009; Kevin Lau, "HP TouchSmart: More than just a PC". *GetConnected.com*, acesso em 31 jul. 2009; Eric Lai, "HP aims TouchSmart desktop PC at businesses". *Computerworld*, 1 ago. 2009; Ina Fried, "Touch in Windows 7: Just for show?". *CNET News*, 1 jul. 2009; e Walter Mossberg, "Multitouch interface is starting to spread among new devices". *The Wall Street Journal*, 31 jan. 2008.

PERGUNTAS SOBRE O ESTUDO DE CASO

1. Que problemas são resolvidos com a tecnologia multitoque?
2. Quais as vantagens e desvantagens das interfaces multitoque? Quão úteis elas podem ser? Explique.
3. Descreva três aplicações de negócios que se beneficiariam de uma interface multitoque.
4. Quais questões humanas, empresariais e tecnológicas devem ser consideradas se você ou sua empresa considerasse adotar sistemas e computadores com interfaces multitoque?

Tabela 4.2 Principais sistemas operacionais de PCs e servidores

Sistema operacional	Características
Windows 7	Sistema operacional Windows mais recente para usuários finais, apresenta melhorias em itens como usabilidade, barra de tarefas, desempenho e segurança, bem como suporte para interfaces multitoque.
Windows Vista	Sistema operacional para PCs poderosos, com versões domésticas e corporativas. Inclui pesquisa no desktop, ferramentas multimídia, melhorias na segurança se comparado às versões anteriores do Windows, e sincronização com dispositivos móveis, câmeras e serviços de Internet.
Windows Server 2008	Versão mais recente do sistema operacional Windows para servidores.
UNIX	Usado em PCs, estações de trabalho e servidores de rede. Suporta multitarefa, processamento multiusuário e rede. Pode ser instalado em diferentes modelos de hardware.
Linux	Sistema operacional com código aberto, alternativa confiável aos sistemas operacionais UNIX e Windows. Roda em diferentes tipos de hardware e pode ser modificado por desenvolvedores de software.
Mac OS X	Sistema operacional para o Macintosh, estável e confiável, com recursos de busca poderosos e que suporta processamento de vídeo e imagens, além de oferecer elegante interface ao usuário. A versão mais recente é a Snow Leopard. O sistema operacional iOS/iPhone/iPad é derivado do OS X.

A família Windows de sistemas operacionais, produzida pela Microsoft, conta com versões cliente e servidor e uma interface gráfica de usuário simplificada. Os sistemas Windows podem executar inúmeras tarefas de programação simultaneamente e têm poderosos recursos de rede, incluindo a capacidade de acessar informações na Internet. O **Windows 7** é sua versão mais recente. Suas melhorias em relação ao **Windows Vista** e ao antigo **Windows XP** incluem aperfeiçoamento da usabilidade, desempenho mais rápido, nova barra de tarefas, suporte para interfaces multitoque e melhorias adicionais na segurança. Existem versões para usuários domésticos, de pequenas empresas e corporativos. O Windows 7 acrescentou uma versão chamada Starter, aos pequenos notebooks e netbooks, nos quais a capacidade de processamento e a memória são limitadas para uso de versão com maior número de recursos.

Os sistemas operacionais Windows para servidores de rede proporcionam funções de gestão de rede, inclusive ferramentas para criar e operar sites e outros serviços de Internet. O **Windows Server 2008** tem versões para grandes, médias e pequenas empresas, e para organizações com vastos centros computacionais e necessidades de processamento.

O **UNIX** é um sistema operacional, multiusuário e multitarefa desenvolvido pelo Bell Laboratories em 1969 para auxiliar o compartilhamento de dados por pesquisadores científicos. Foi projetado para interligar diversas máquinas e proporciona grande capacidade de suporte para comunicações e redes. É frequentemente utilizado em estações de trabalho e servidores, oferecendo confiabilidade e capacidade de ampliação para rodar em sistemas maiores ou **servidores multiusuários** (*high-end*). Pode rodar em diversos tipos de computador e ser facilmente configurado. Programas aplicativos que rodam em UNIX podem ser exportados de um computador a outro com poucas modificações. Foram desenvolvidas interfaces gráficas de usuário para esse sistema. O UNIX também apresenta alguns problemas de segurança, pois permite que múltiplas tarefas e usuários acessem o mesmo arquivo simultaneamente. Além disso, os fornecedores do UNIX desenvolveram versões diferentes, porém incompatíveis, do programa, limitando assim a portabilidade do software.

O **Linux** é um sistema operacional semelhante ao UNIX, que pode ser baixado da Internet gratuitamente ou comprado mediante pequena taxa, de empresas que fornecem ferramentas adicionais para o software. É grátis, confiável, de projeto compacto e habilitado a rodar em muitas plataformas diferentes de hardware, dentre elas servidores, computadores de mão e eletroeletrônicos de consumo.

Nos últimos anos, o Linux ganhou popularidade como alternativa robusta e de baixo custo ao UNIX e à família Windows. Por exemplo, a E*Trade Financial economiza 13 milhões de dólares por ano com a melhoria no desempenho dos computadores através da execução do Linux em uma série de servidores IBM de baixo custo, em vez de caros servidores Sun Microsystems executando a versão proprietária do UNIX da Sun.

O Linux tem papel decisivo no *back-office*, sendo utilizado em servidores da Web e redes locais. Atualmente, o 'sistema do pinguim' detém 25 por cento do mercado norte-americano de servidores. Ao mesmo tempo, seu uso em computadores de mesa está crescendo a passos largos. IBM, HP, Intel, Dell e Sun fizeram do Linux um atrativo central em suas ofertas a clientes corporativos, enquanto grandes fornecedores de software estão começando a oferecer versões de seus produtos que podem rodar no Linux. Tanto a IBM quanto a Sun oferecem aplicativos gratuitos, ou de baixo custo, para o Linux.

O Linux é um exemplo de **software de código aberto** que permite acesso livre a seu código de programa, de modo que todos os usuários de computadores possam modificá-lo para corrigir erros ou fazer melhorias. Softwares de código aberto como o Linux não pertencem a nenhuma empresa ou indivíduo. Quem gerencia e modifica o software é uma rede global de programadores e usuários, que geralmente não são pagos para fazê-lo. Por definição, os softwares de código aberto não estão restritos a um sistema operacional ou a uma tecnologia de hardware específica, embora a maioria deles atualmente se baseie no Linux ou no UNIX.

Além desses sistemas operacionais, estão surgindo novos sistemas para dispositivos móveis e computadores conectados à nuvem. O *Chrome OS*, da Google, serve como um sistema operacional mais leve para usuários que executam a maior parte de suas tarefas computacionais na Internet e funciona em computadores que vão desde netbooks até desktops. Ele possui interface minimalista para tirar proveito da Web e da computação em nuvem. O *Android* é um sistema operacional móvel inicialmente desenvolvido pela Google e, mais tarde, pela Open Handset Alliance, como plataforma flexível e atualizável para dispositivos móveis. Ele pode, eventualmente, ser usado em pequenos computadores. A Microsoft lançou um sistema operacional para a nuvem chamado *Windows Azure*, que atende a seus serviços e sua plataforma na nuvem.

Software aplicativo e ferramentas de produtividade para PCs

Atualmente, as empresas têm acesso a uma gama de ferramentas para desenvolver seus softwares aplicativos. Entre elas estão as linguagens de programação tradicional, as linguagens de quarta geração, os pacotes de software aplicativos e ferramentas de produtividade para PCs, os softwares para desenvolvimento de aplicativos de Internet e os softwares para integração empresarial. É importante saber quais ferramentas de software e linguagens de programação são mais adequadas ao trabalho que a empresa pretende executar.

Linguagem de programação de aplicativos para empresas

Para aplicações empresariais, as mais importantes linguagens de programação são, tradicionalmente, C, C++, Visual Basic e COBOL. Desenvolvida no início da década de 1970, C é uma linguagem poderosa e eficiente que combina a portabilidade de máquina com controle rígido e uso eficiente de recursos computacionais. É usada basicamente por programadores profissionais na criação de sistemas operacionais e software aplicativo, especialmente para PCs. C++ é uma visão mais recente da linguagem C, com todos os seus recursos mais a possibilidade de trabalhar com objetos de software. Diferentemente dos programas tradicionais, que separam dados das ações a serem feitas com eles, um **objeto** de software combina dados e procedimentos. O Capítulo 11 descreve em detalhes o desenvolvimento de software orientado a objetos. O **Visual Basic** é uma ferramenta de programação amplamente utilizada para criar aplicações que rodam no Microsoft Windows. Uma **linguagem de programação visual** permite ao usuário manipular elementos gráficos ou ícones para criar programas. O **COBOL** (*COmmon Business Oriented Language*) foi desenvolvido no início da década de 1960 para processar grandes arquivos de dados com caracteres alfanuméricos (que misturam dados alfabéticos e numéricos) e para produzir relatórios corporativos. Atualmente, a COBOL é encontrada em grandes sistemas empresariais legados.

Linguagens de quarta geração


Linguagens de quarta geração consistem em uma variedade de ferramentas de software que habilitam usuários finais a desenvolver softwares aplicativos com o mínimo de assistência técnica (ou nenhuma) ou que proporcionem ganhos de produtividade aos programadores profissionais. Tendem a ser não orientadas ou menos orientadas a procedimentos do que as linguagens convencionais. As linguagens orientadas a procedimentos requerem a especificação de sequências de passos ou procedimentos que dizem ao computador o que e como fazer. As linguagens não orientadas a procedimentos apenas especificam o que precisa ser executado, em vez de dar detalhes sobre a execução da tarefa. Algumas das linguagens não orientadas a procedimentos são as **linguagens naturais**, que habilitam os usuários a se comunicar com o computador usando comandos em linguagem comum, semelhante à fala humana.

A Tabela 4.3 mostra que há seis tipos de linguagens de quarta geração: ferramentas de software de PCs, linguagens de consulta, geradores de relatórios, linguagens gráficas, geradores de aplicações e pacotes de softwares aplicativos. As ferramentas apresentadas na tabela estão classificadas por ordem de facilidade de utilização por usuários finais que não sejam programadores. O mais provável é que os usuários finais trabalhem com ferramentas de software de PCs e **linguagens de consulta**, que dão respostas imediatas on-line a requisições de informação não predefinidas, tais como: “quais são os representantes de vendas que têm o melhor desempenho?”. As linguagens de consulta muitas vezes estão vinculadas a softwares de gestão de dados e a sistemas de gestão de banco de dados (veja o Capítulo 5).

Pacotes de software e ferramentas de produtividade para PCs

Grande parte dos softwares usados nas empresas hoje não é programada sob medida; consiste, isso sim, em pacotes de software aplicativo e ferramentas de produtividade para PCs. Um **pacote de software** é um conjunto de programas escritos e codificados previamente, disponíveis no mercado, que eliminam a necessidade de indivíduos e empresas

Tabela 4.3 Categorias de linguagem de quarta geração

<u>Ferramenta de linguagem de quarta geração</u>	<u>Descrição</u>	<u>Exemplo</u>	<u>Orientada ao usuário final</u>
Ferramentas de software para PC	Pacotes de software aplicativo de uso geral para PCs.	Word Perfect Microsoft Access	
Linguagens de consulta	Linguagens para extrair dados armazenados em arquivos ou bancos de dados. Suportam requisições de informação não predefinidas.	SQL	
Geradores de relatórios	Extraem dados de arquivos ou bancos de dados para criar relatórios customizados sob uma grande variedade de formatos não produzidos por sistemas de informação. Geralmente proporcionam maior controle sobre a maneira como os dados são formatados, organizados e apresentados do que as linguagens de consulta.	Crystal Reports	
Linguagens gráficas	Extraem dados de arquivos ou bancos de dados e os apresentam sob o formato de gráficos. Alguns softwares geradores de gráficos também podem executar operações aritméticas ou lógicas com os dados.	SAS/Graph Systat	
Geradores de aplicações	Contêm módulos pré-programados que podem gerar aplicações completas, incluindo sites, conferindo grande velocidade ao desenvolvimento. O usuário pode especificar o que precisa ser feito e o gerador de aplicação criará o código de programa apropriado para entrada, validação, atualização e apresentação.	WebFOCUS QuickBase	
Pacotes de softwares aplicativos	Programas de software vendidos ou arrendados por fornecedores que eliminam a necessidade de software personalizado.	Oracle PeopleSoft HCM mySAP ERP	

escreverem programas de software para determinadas funções. Existem pacotes de software de sistema, mas a maioria é de softwares aplicativos.

A instalação e o suporte de pacotes de software que rodam em mainframes e computadores de maior porte comumente requerem programadores profissionais. Pacotes de produtividade elaborados para PCs, como editores de texto, planilhas eletrônicas, gestão de dados, recursos gráficos de apresentação e navegadores da Web são as ferramentas de software mais amplamente utilizadas por empresas e usuários individuais.

Software de processamento de texto Se você trabalha ou estuda, provavelmente usa algum software de processamento de texto todos os dias. Um **software de processamento de textos** armazena textos eletronicamente como arquivos de computador, em vez de arquivos em papel. Permite que o usuário faça alterações eletronicamente em um documento na memória, eliminando a necessidade de digitar novamente a página inteira para incorporar as correções. Tem opções de formatação que permitem mudar o espaçamento entre as linhas, o espaçamento entre as margens, o tamanho das letras e a largura das colunas. O Microsoft Word e o WordPerfect são pacotes de processadores de texto muito comuns.

A maioria dos softwares de processamento de texto dispõe de recursos avançados que automatizam outras tarefas de redação: verificadores ortográficos, verificadores de estilo (analisam gramática e pontuação), dicionários de sinônimos e outros significados (tesauros), programas de mala direta (que vinculam cartas ou outros documentos a nomes e endereços de uma lista de endereçamento), e recursos para criação de páginas da Web e acesso a elas.

Empresas que precisam criar folhetos, manuais ou livros provavelmente utilizarão softwares de editoração eletrônica. Esse tipo de software permite maior controle sobre a distribuição de texto, os recursos gráficos e as fotos em uma página (*layout*) do que um software de edição de texto. O Adobe InDesign e o QuarkXpress são dois pacotes de editoração eletrônica muito conhecidos.

Software de planilha eletrônica Fornece versões computadorizadas de ferramentas tradicionais de modelagem financeira, como o livro de registros contábeis com lápis e calculadora. A **planilha eletrônica** é organizada como uma grade de colunas e linhas. Os recursos que ela oferece ficam evidentes quando trocamos um valor ou uma série de valores; por exemplo, os outros valores relacionados são automaticamente recalculados.

Com frequência, as planilhas eletrônicas são usadas em aplicações que exigem modelagem e análise de simulação. Após o usuário ter montado um conjunto de relações matemáticas, a planilha pode ser recalculada instantaneamente usando-se um conjunto diferente de suposições. Várias alternativas podem ser avaliadas com facilidade trocando-se um ou dois dados, sem que seja necessário digitar novamente todo o restante da planilha. Os pacotes de planilhas eletrônicas incluem funções gráficas com dados sob a forma de gráficos de linha, diagramas de barras ou gráficos de setores circulares (pizza). O pacote de planilha eletrônica mais conhecido é o Microsoft Excel. As versões mais recentes desse software podem ler e escrever arquivos da Web. A Figura 4.8 ilustra a planilha resultante de uma análise de ponto de equilíbrio e o gráfico correspondente.

Softwares gerenciadores de bancos de dados Embora os programas de planilhas eletrônicas sejam poderosas ferramentas para a manipulação de dados quantitativos, o software **gerenciador de bancos de dados** é mais adequado para criar e manipular listas e combinar informações extraídas de arquivos diferentes. Os pacotes de gestão de bancos de dados para PCs em geral têm recursos de programação e menus intuitivos que tornam possível, inclusive para leigos, construir pequenos sistemas de informação.

Normalmente, os softwares gerenciadores de dados facilitam a criação de arquivos e bancos de dados e o armazenamento, a modificação e a manipulação de dados para relatórios e consultas. Dentre os softwares gerenciadores de bancos de dados muito conhecidos para computadores pessoais está o Microsoft Access, aperfeiçoado para editar dados na Web. Programas de gestão de bancos de dados são discutidos em detalhes no Capítulo 5.

Recursos gráficos de apresentação O software de **recursos gráficos de apresentação** permite que o usuário crie apresentações gráficas de qualidade profissional. Esse software

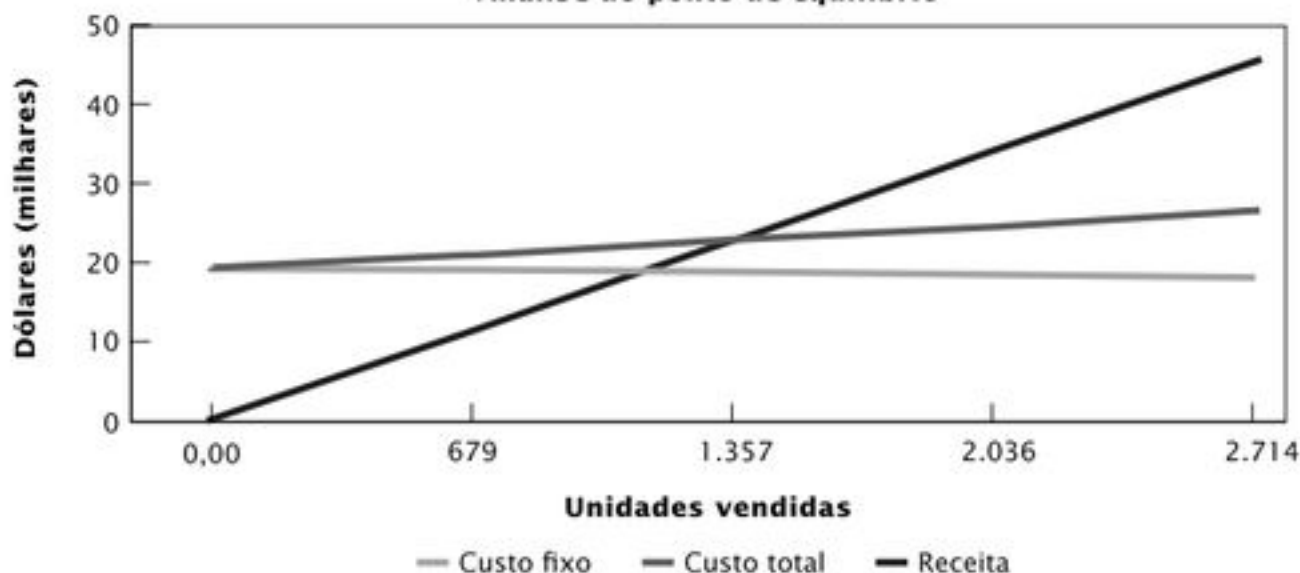
Figura 4.8**Software de planilha eletrônica**

Este software organiza dados em colunas e linhas para análise e manipulação. Os softwares de planilha mais modernos oferecem recursos gráficos para apresentação visual clara dos dados. Esta amostra de análise de ponto de equilíbrio é representada como números sobre a planilha e também sob a forma de um gráfico de linha, que facilita a interpretação dos dados.

Custo fixo total	19.000,00
Custo variável por unidade	3,00
Preço médio de venda	17,00
Margem de contribuição	14,00
Ponto de equilíbrio	1.357

Gravatas modelo tradicional Demonstração de resultados pro forma

Unidades vendidas	0,00	679	1.357	2.036	2.714
Receita	0	11.536	23.071	34.607	46.143
Custo fixo	19.000	19.000	19.000	19.000	19.000
Custo variável	0	2.036	4.071	6.107	8.143
Custo total	19.000	21.036	23.071	25.107	27.143
Lucros/perdas	(19.000)	(9.500)	0	9.500	19.000

**Gravatas modelo tradicional
Análise do ponto de equilíbrio**

pode converter dados numéricos em diagramas e outros tipos de gráficos e incluir recursos de apresentação multimídia de som, animação, fotos e clipes de vídeo. Os principais pacotes de recursos gráficos de apresentação incluem recursos para apresentar slides gerados por computador e para transferir conteúdo para a Web. O Microsoft PowerPoint e o Lotus Freelance Graphics são pacotes de recursos gráficos de apresentação bastante conhecidos.

Suítes de software Em geral, os principais aplicativos de produtividade estão agrupados em suítes de software. O Microsoft Office é um exemplo. Existem diferentes versões para usuários corporativos e domésticos, mas as principais ferramentas incluem o software processador de texto Word, o software de planilha eletrônica Excel, o software de banco de dados Access, o software de apresentação gráfica PowerPoint e o Outlook, um conjunto de ferramentas para e-mail, agendamento e gestão de contatos. O Microsoft Office 2010 é a versão mais atual dessa suíte. A Microsoft inclui uma versão da Web denominada Office Web Apps, disponível para usuários e empresas através do pagamento de uma taxa de assinatura ou gratuitamente. O Office Web Apps distribui versões on-line mais leves de ferramentas da suíte, como Word, Excel, PowerPoint e One Note através de navegadores-padrão para PCs, telefones móveis e outros dispositivos disponíveis para a rede. A Microsoft armazena os dados dos usuários dessas aplicações on-line e permite certo grau de compartilhamento de documentos e arquivos.

Existem suítes de ferramentas de produtividade de baixo custo competindo com o Microsoft Office, como o OpenOffice (que pode ser baixado gratuitamente da Internet) e o StarOffice (adquirido por 34,95 dólares), ambos da Sun Microsystems. Entretanto, a ameaça real ao Microsoft Office está vindo da nuvem. Versões na internet para aplicativos de produtividade estão se tornando populares devido à conveniência, à flexibilidade e ao baixo custo. Existem mais de 2 milhões de empresas de todos os tamanhos utilizando o Google Apps, apresentado no Capítulo 2 em nossa discussão sobre ferramentas colaborativas. Essas suítes on-line incluem ferramentas para processamento de textos, planilhas

eletrônicas, apresentações, gestão de contatos, mensagens e e-mail, e estão disponíveis gratuitamente ou para as empresas em uma versão Premier, com mais recursos, a 50 dólares.

Navegadores da Web Os navegadores da Web são ferramentas de software fáceis de usar, projetadas para apresentar páginas da rede e acessar a Web e outros recursos de Internet. Os navegadores podem exibir ou apresentar informações gráficas em áudio e vídeo, assim como em texto tradicional, e permitem clicar (ou tocar) em botões ou palavras iluminadas na tela para estabelecer links com sites relacionados. Os navegadores da Web transformaram-se na interface primária para acessar a Internet ou para usar sistemas em rede com tecnologia de Internet. Os principais navegadores existentes no mercado são o Internet Explorer, da Microsoft; o Mozilla Firefox; o Safari, da Apple; e o Google Chrome. Os dispositivos móveis possuem seus próprios navegadores especializados.

Softwares para a Web: Java, Ajax e HTML

As empresas usam ferramentas de software especiais para desenvolver sites e aplicativos que rodam na Web. Java e Ajax são usadas para o desenvolvimento de aplicativos que rodam na Web, enquanto a HTML é usada para a criação de páginas da Web.

Java

Java é uma linguagem de programação orientada a objeto e que não está vinculada a um processador ou sistema operacional específico, e tornou-se o principal ambiente de programação interativo para a Web. A linguagem Java permite que os usuários manipulem os dados em sistemas em rede usando navegadores da Web, reduzindo a necessidade de escrever software especializado. No âmbito empresarial, é usada para aplicações mais complexas de E-commerce e E-business, que exigem comunicação com os sistemas de processamento de transações de *back-end* da organização.

Praticamente todos os softwares de navegador da Web têm uma plataforma Java embutida. A plataforma Java migrou para celulares, automóveis, players de música, videogames e, finalmente, sistemas interativos de televisão a cabo, nos quais são usados para disponibilizar conteúdo interativo.

O software Java é projetado para rodar em qualquer computador ou dispositivo de computação, independentemente do microprocessador ou sistema operacional específico utilizado. A linguagem consegue realizar essa façanha utilizando uma máquina virtual Java construída para cada tipo de computador e sistema operacional. Um PC Macintosh, um PC IBM rodando sobre Windows, um servidor Sun rodando em UNIX e até mesmo um telefone celular ou um PDA podem compartilhar a mesma aplicação Java, reduzindo os custos do desenvolvimento de software e criando a mesma experiência para o usuário, independentemente do tipo de computador com o qual ele está trabalhando.

Em ambientes de rede, como a Internet, a linguagem Java é utilizada para criar programas em miniatura chamados '*applets*', projetados para residir em servidores de redes centralizadas. A rede apresenta apenas os *applets* exigidos por uma função específica. Quando há *applets* Java residindo em uma rede, o usuário pode descarregar apenas as funções de software e os dados de que precisa para executar determinada tarefa, tal como analisar a receita produzida por um território de vendas. O usuário não precisa manter extensos programas de software nem arquivos de dados em sua máquina.

Ajax

Você já preencheu um formulário na Web, cometeu algum erro e precisou recomeçar após uma longa espera para que uma nova página de formulário de pedido aparecesse em sua tela? Ou visitou um site de mapas, clicou uma vez na seta para o Norte e esperou algum tempo até que uma página inteira fosse recarregada? **Ajax** (*Asynchronous JavaScript and XML*) é outra técnica de desenvolvimento para criação de aplicações da Web interativas que evitam todas essas inconveniências.

Ele permite que um cliente e um servidor troquem pequenos dados nos bastidores para que não haja necessidade de recarregar uma página da Web inteira a cada vez que um usuário

faz uma solicitação. Portanto, se clicar em Norte em um site de mapas, como o Google Maps, o servidor transfere somente a parte da aplicação a ser alterada, sem a necessidade de espera por um novo mapa inteiro. Em aplicações de mapas, também é possível arrastar e mover os mapas em qualquer direção sem forçar a recarga da página toda. Programas desenvolvidos em Ajax utilizam programas em JavaScript baixados para o cliente a fim de manter uma conversa quase contínua com o servidor utilizado, tornando a experiência do usuário mais realista.

Hypertext Markup Language (HTML)

A **Hypertext Markup Language (HTML)** é uma linguagem de descrição de página que especifica como texto, recursos gráficos, vídeo e som serão organizados em uma página da Web para criar links dinâmicos com outras páginas e objetos. Utilizando esses links, basta que o usuário aponte para uma palavra-chave ou um recurso gráfico iluminados e clique sobre eles, que será transportado imediatamente para outro documento. A Tabela 4.4 exemplifica algumas declarações da linguagem HTML.

Os programas HTML podem ser escritos conforme especificação, mas também podem ser criados usando os recursos de criação e autoria dos navegadores da Web ou de populares editores de texto, planilhas eletrônicas, gestão de dados e pacotes de software de apresentação gráfica. Editores HTML como Adobe Dreamweaver são os programas de ferramentas de autoria HTML mais poderosos para criação de páginas da Web.

Serviços da Web

Por **serviços da Web (Web services)**, entenda-se um conjunto de componentes de software vagamente relacionados que trocam informações entre si usando linguagens e padrões de comunicação universais da Web. Com eles, é possível trocar informações entre dois sistemas, independentemente dos sistemas operacionais ou das linguagens de programação nos quais eles estiverem baseados. Os serviços da Web podem ser usados para desenvolver aplicativos da Internet, de padrão aberto, que conectam os sistemas de duas organizações diferentes e, também, para criar aplicativos que integram sistemas díspares dentro de uma mesma empresa. Os serviços da Web não estão relacionados a nenhuma linguagem de programação ou sistema operacional específico, e diferentes aplicativos podem usá-los para se comunicar de maneira padronizada, sem a necessidade de gastar tempo com codificações.

A tecnologia básica para os serviços da Web é a **XML**, abreviatura de **eXtensible Markup Language**. Essa linguagem foi desenvolvida em 1996 pelo World Wide Web Consortium (W3C, o órgão internacional que supervisiona o desenvolvimento da Web) para ser uma linguagem de markup mais flexível e poderosa que a HTML para páginas da rede. Enquanto a HTML apenas determina como textos e imagens devem ser apresentados em um documento Web, a XML descreve o que significam os dados apresentados nesses documentos, de modo que possam ser utilizados em programas de computador. Na XML, um número não é simplesmente um número; a tag do XML especifica se o número representa um preço, uma data ou um código de endereçamento postal (CEP). A Tabela 4.5 exemplifica algumas declarações de XML.

Ao colocar o significado de elementos selecionados do conteúdo dos documentos entre tags, a XML faz com que os computadores possam manipular e interpretar automaticamente seus dados e executar operações com eles sem intervenção humana. Navegadores da Web e programas de computador como os softwares de processamento de pedidos e

Tabela 4.4 Exemplos de HTML

Em português	HTML
Subcompacto	<TITLE>Automóvel</TITLE>
4 passageiros	4 passageiros
\$16.800	\$16.800

Tabela 4.5 Exemplos de XML

Em português	XML
Subcompacto	<TIPO DO AUTOMÓVEL>= "Subcompacto">
4 passageiros	<UNIDADE DE PASSAGEIRO="PASS">4</PASSAGEIROS>
\$ 16.800	<MOEDA DO PREÇO="USD"> \$ 16.800</PREÇO>

ERP podem seguir regras programadas para aplicar e apresentar dados. A XML provê um formato-padrão para a permuta de dados, permitindo que os serviços da Web transmitam dados de um processo para outro.

Os serviços Web comunicam-se por meio de mensagens XML codificadas em protocolos-padrão da Web. O protocolo *SOAP*, acrônimo de *Simple Object Access Protocol* (protocolo de acesso de objeto simples), é um conjunto de regras para estruturar mensagens que permite que as aplicações passem dados e instruções umas para as outras. A *WSDL*, que significa *Web Services Description Language* (linguagem de descrição de serviços Web), é uma estrutura comum para descrever as tarefas realizadas por um serviço Web, bem como os comandos e os dados que ele aceitará, de maneira que possa ser usado por outras aplicações. O *UDDI*, acrônimo de *Universal Description Discovery and Integration* (descrição universal de descoberta e integração), permite que um serviço Web seja listado em um diretório de serviços de modo a ser facilmente localizado. As empresas podem descobrir e encontrar serviços da Web nesse diretório da mesma maneira que fariam nas Páginas Amarelas ou em uma lista telefônica. Usando esses padrões e protocolos, uma aplicação pode se conectar livremente com outras aplicações sem programação customizada para cada uma das aplicações diferentes com a qual deseja comunicar-se. Todas compartilham os mesmos padrões.

O conjunto de serviços Web usados para desenvolver os sistemas de software de uma empresa constitui o que chamamos de arquitetura orientada a serviços. Uma **arquitetura orientada a serviços (SOA)** é um conjunto autossuficiente de serviços que se comunicam entre si a fim de criar um aplicativo de software útil à empresa. Para realizar as tarefas organizacionais, uma sequência desses serviços é executada. Se necessário, os desenvolvedores de software podem utilizar tais serviços também em outras combinações, para montar outros aplicativos.

Praticamente todos os grandes fornecedores de software, como a IBM, a Microsoft, a Sun e a HP, oferecem ferramentas e plataformas inteiras para desenvolver e integrar aplicativos de software usando serviços Web. A IBM inclui ferramentas de serviços Web na WebSphere, sua plataforma de software para E-business, e a Microsoft incorporou ferramentas parecidas na plataforma Microsoft.NET.

Os sistemas da locadora de veículos Dollar Rent-A-Car usam serviços Web para conectar seu sistema de reservas on-line ao site da Southwest Airlines. Embora os sistemas das duas empresas estejam baseados em plataformas tecnológicas diferentes, um cliente que reserve um voo no Southwest.com pode locar um carro da Dollar sem sair do site da companhia aérea. Em vez de lutar para que o sistema de reservas da Dollar compartilhasse dados com os sistemas de informação da Southwest, a Dollar usou a tecnologia de serviços Web do Microsoft.NET para fazer a intermediação. As reservas da Southwest são traduzidas em protocolos de serviços Web, que, por sua vez, são passados para um formato que os computadores da Dollar conseguem entender.

Outras autolocatoras já haviam conectado seus sistemas de informação a sites de companhias aéreas. Mas, sem os serviços Web, essas conexões precisavam ser desenvolvidas uma a uma. Os serviços Web oferecem código padronizado no qual os computadores da Dollar podem 'conversar' com os sistemas de informação das outras empresas, sem a necessidade de construir links especiais para cada um. A Dollar está expandindo seu uso dos serviços da Web para conectar-se diretamente a um pequeno operador turístico e a um grande sistema de reservas de viagem, assim como a um site de **Web sem fio** para

celulares e PDAs. Para tanto, não é necessário escrever novos códigos de software para o sistema de informação de cada novo parceiro nem para cada novo dispositivo sem fio (veja a Figura 4.9).

Tendências de software

Hoje, existem muitas fontes para obtenção de software e muito mais recursos para que os usuários criem suas próprias aplicações de software customizadas. O uso crescente de software de código aberto e de ferramentas de software baseadas na nuvem exemplifica essa tendência.

Software de código aberto

A tendência de software mais influente é o movimento pelo software de código aberto. Como já observado, o software de código aberto é desenvolvido por uma comunidade de programadores ao redor do mundo que disponibilizam seus programas aos usuários sob um dos vários esquemas de licença disponíveis. Basicamente, os usuários podem utilizar o software tal qual se apresenta, modificá-lo como quiserem e mesmo incluí-lo em aplicações de software com fins lucrativos.

O movimento pelo código aberto teve início em 1983 (quando foi chamado 'software hippie'), mas desde então se desenvolveu e assumiu um papel importante na infraestrutura empresarial de computação, com o surgimento de programas como Linux e Apache, o software servidor Web mais largamente utilizado. Atualmente é possível encontrar milhares de programas de código aberto para realizar tarefas que vão desde carrinhos de compra de E-commerce e liberação de fundos até a gestão da força de vendas. Algumas aplicações da computação em nuvem descritas neste capítulo são baseadas em software de código aberto, como o navegador Google Chrome.

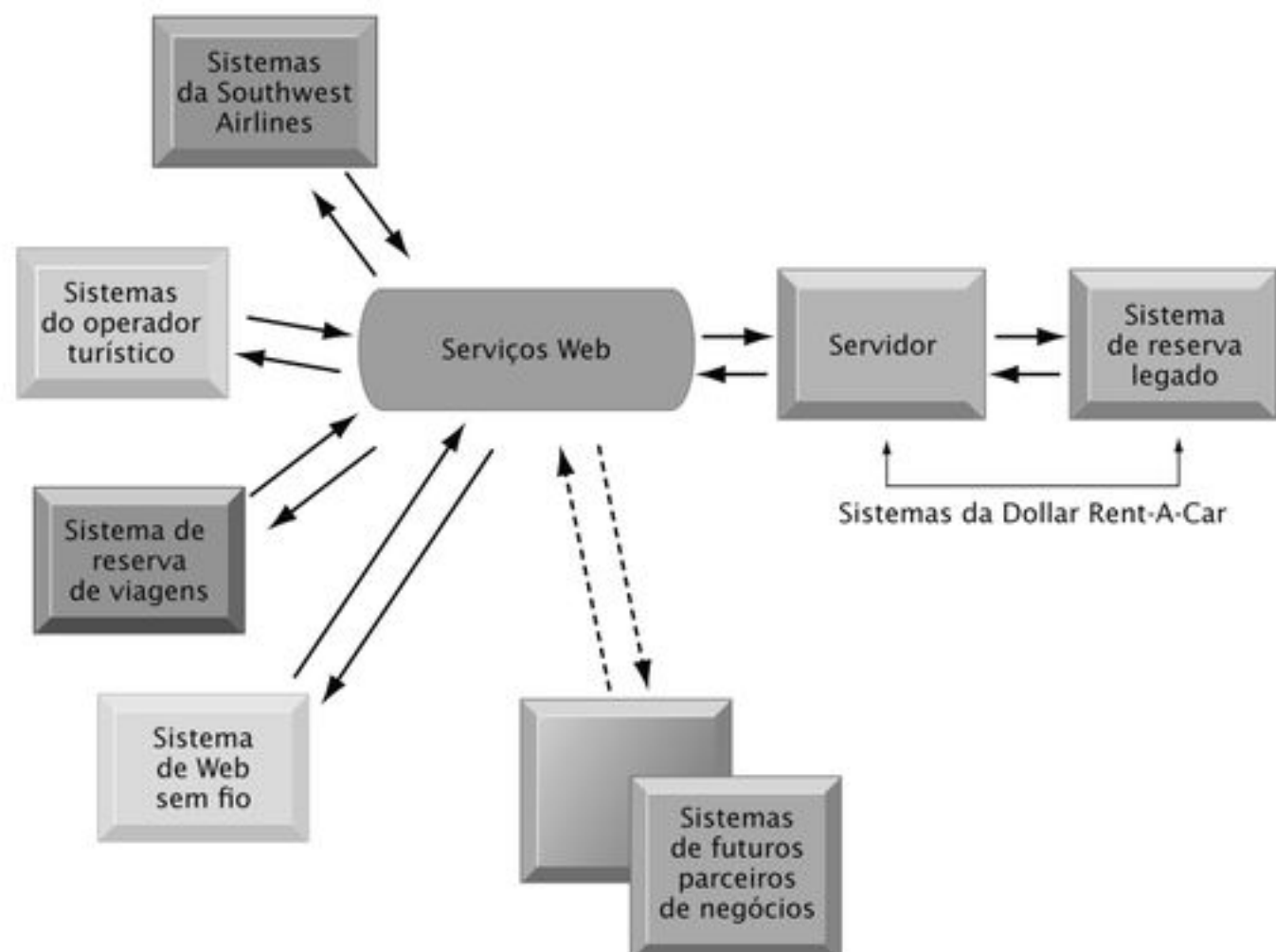
Ferramentas e serviços baseados na nuvem

No passado, softwares como o Microsoft Word ou o Adobe Ilustrator vinham em uma caixa e eram projetados para rodar em uma única máquina. Atualmente, é mais provável que se faça download do software do site do fornecedor para o computador ou, cada vez mais, que se utilize o software como um serviço da nuvem disponibilizado pela Internet.

Figura 4.9

Como a Dollar Rent-A-Car usa serviços Web

Com os serviços Web, a Dollar Rent-A-Car conta com uma camada de software intermediária padronizada para 'conversar' com os sistemas de informação de outras empresas. A Dollar pode usar esse conjunto de serviços Web para conectar-se aos sistemas de informação de outras organizações sem ter de desenvolver um link separado para cada um.



O software baseado na nuvem e os dados por ele utilizados estão hospedados em servidores poderosos em grandes centros de dados, e podem ser acessados por qualquer pessoa com uma conexão à Internet e um navegador-padrão da Web. As inúmeras aplicações baseadas na Web da Google, descritas no início deste capítulo e no Capítulo 2, são um exemplo. Além do Office Web Apps, a Microsoft oferece outros serviços de software na nuvem, como o *Business Productivity Online Standard Suite* para mensagens e colaboração.

Mashups e widgets O software utilizado tanto para tarefas pessoais quanto para tarefas empresariais pode ser composto por programas grandes e independentes, ou pode ser formado por componentes intercambiáveis que se integram livremente a outras aplicações na Internet. Usuários individuais e empresas inteiras combinam e adaptam esses componentes de software para criar sua própria aplicação customizada e compartilhar informações com os outros. As aplicações de software resultantes são chamadas *mashups*. Você já executou um *mashup* se configurou seu perfil no Facebook ou seu blog com o recurso de exibição de vídeos e apresentações.

A ideia é partir de diferentes fontes e produzir um trabalho novo, superior à soma das partes. Integrantes de um movimento chamado Web 2.0 (veja o Capítulo 6) e com o espírito dos *mashups* musicais, os *mashups* da Web combinam os recursos de dois ou mais aplicativos on-line para criar uma espécie de híbrido, que oferece mais valor ao cliente do que as fontes originais sozinhas. O Faceforce, por exemplo, integra as informações do perfil do Facebook aos dados do Salesforce em tempo real, oferecendo uma visão imediata completa de seus clientes reais e potenciais, e dos sócios, e um local único para visualização e gestão de todas as informações.

Uma área de grande inovação é o *mashup* de software para mapeamento e imagem por satélite com conteúdo local. Google, Yahoo! e Microsoft agora oferecem ferramentas para permitir que outras aplicações utilizem informações de seus mapas e imagens de satélite com relativamente pouca programação. O ZipRealty, por exemplo, utiliza o Google Maps e os dados fornecidos pela comunidade imobiliária on-line Zillow.com para exibir uma listagem completa de imobiliárias do Serviço de Listagem Múltipla (SLM) para qualquer código postal definido pelo usuário no Estados Unidos. O BidNearBy utiliza o Google Maps e dados da Craigslist do eBay para pesquisar leilões locais e listas de classificados e exibir sua localização em forma de mapa.

Os pequenos pedaços de código de software que permitem ao usuário incluir conteúdo de um site em outra página ou outro site são chamados *widgets*. Eles são pequenos programas de software que podem ser acrescentados a páginas da Web ou instalados no desktop (ou no dispositivo digital móvel) para oferecer funcionalidade adicional. O *widget* Amo Cinema do perfil do Facebook, por exemplo, leva os usuários a um local no qual podem listar os filmes que já viram e associar a eles uma nota e um comentário, visualizar as notas e comentários de seus amigos e o que está em cartaz no cinema. O *widget* Atom, do iPhone, distribui novas notícias do Google para o iPhone, e existem *widgets* para exibição de atualizações do Twitter em seu blog.

Os *widgets* da Web funcionam dentro de uma página ou blog. Os *widgets* para desktop integram conteúdo de uma fonte externa ao desktop do usuário para oferecer serviços como calculadora e dicionário, ou para exibir condições meteorológicas atuais. Yahoo! Weather, Dashboard TV, da Apple, e Google Desktop Gadgets são exemplos de *widget* para desktop.

Os *widgets* também podem oferecer uma vitrine para propaganda e venda de produtos e serviços. Amazon e Walmart possuem *widgets* para barras de ferramentas que permitem que os usuários pesquisem suas lojas virtuais enquanto estiverem em outra página da Web. Os *widgets* tornaram-se tão poderosos e úteis que Facebook e Google lançaram programas para atrair desenvolvedores de *widgets* para seus sites.

Software como serviço (SaaS – software as a service). Além das ferramentas gratuitas ou de baixo custo para indivíduos e empresas oferecidas por empresas como Google, Microsoft ou Yahoo!, softwares integrados e outras funções de negócios complexas estão disponíveis como serviços disponibilizados pelos principais fornecedores de software. Em vez de comprar e instalar programas, empresas alugam as mesmas funções desses serviços,

com usuários pagando uma assinatura ou uma taxa por transação. Serviços para disponibilização de softwares e acesso remoto aos mesmos na forma de serviço Web são agora chamados de **software como serviço (SaaS)**.

Um exemplo é o Salesforce.com, que oferece serviços de software sob demanda para gestão do relacionamento com o cliente, incluindo automação da força de vendas, gestão do relacionamento com o parceiro, marketing e serviços aos clientes. Ele inclui ferramentas para customização, integrando seu software com outras aplicações corporativas e criando novas aplicações. Você vai saber mais sobre o Salesforce.com no estudo de caso ao final do capítulo.

Administração dos recursos de hardware e software

A seleção e a utilização de tecnologia de hardware e software podem causar profundo impacto sobre o desempenho da empresa. Vamos abordar agora as questões mais importantes que você enfrentará ao gerenciar os recursos tecnológicos de hardware e software: planejamento de capacidade e escalabilidade; determinação do custo total de propriedade dos recursos tecnológicos; e decisão se a empresa vai possuir e manter seu próprio hardware, software e outros componentes de infraestrutura, ou se vai arrendá-los de um provedor externo de tecnologia; e gestão de plataformas móveis e de localização de software.

Planejamento de capacidade e escalabilidade

O E-commerce e o E-business estão impingindo novas e intensas demandas sobre a tecnologia de hardware. São necessários recursos de processamento e armazenamento muito maiores para processar e armazenar as transações digitais emergentes que fluem entre diferentes partes da empresa, e entre ela e seus clientes e fornecedores. A utilização simultânea de um site por muitas pessoas acarreta uma pressão muito grande sobre um sistema de computação. O mesmo acontece com a hospedagem de grandes quantidades de páginas da Web intensivas em recursos gráficos ou vídeos.

Atualmente, administradores e especialistas de sistemas de informação precisam dar mais atenção ao planejamento da capacidade e da escalabilidade do hardware do que no passado. **Planejamento da capacidade** é o processo de prever quando um sistema de hardware de computador ficará saturado. Entre os fatores considerados nessa previsão estão o número máximo de usuários que o sistema pode atender ao mesmo tempo, o impacto causado pelas aplicações de software existentes ou futuras, e medições de desempenho como o tempo mínimo de resposta para o processamento de transações empresariais. O planejamento de capacidade assegura que a organização tenha recursos suficientes para suas necessidades correntes e futuras. Por exemplo, a Bolsa de Valores Nasdaq (*Nasdaq Stock Market*) realiza planejamento contínuo de capacidade para identificar picos no volume de transações com ações e garantir suficiente capacidade de computação a fim de administrar grandes ondas, quando o volume de transações estiver muito pesado.

Embora o planejamento de capacidade seja executado por especialistas em sistemas de informação, os dados fornecidos pelos administradores da empresa são essenciais. Eles devem determinar níveis aceitáveis de **tempo de resposta** e disponibilidade de computador para que os sistemas de missão crítica da empresa mantenham o nível de desempenho empresarial esperado. Novas aplicações, fusões, aquisições e alterações no volume de negócios causarão impacto sobre a carga de trabalho do computador e devem ser consideradas no planejamento da capacidade do hardware.

Escalabilidade refere-se à capacidade que um computador, produto ou sistema tem de expandir-se para servir a um número maior de usuários sem sofrer pane. Tanto o E-commerce quanto o E-business exigem infraestruturas de TI escaláveis, que possam crescer com a empresa à medida que aumentarem o tamanho e o número de visitantes do site. As organizações devem certificar-se de possuir recursos suficientes de processamento, armazenagem e rede para administrar volumes crescentes de transações digitais e para disponibilizar tais dados on-line instantaneamente.

Custo total de propriedade de recursos tecnológicos

Se pretende calcular quanto custam o hardware e o software de sua empresa, o preço de compra é apenas a primeira informação a considerar. Também é preciso levar em conta os custos contínuos de administrar atualizações de hardware e software, manutenção, suporte técnico, treinamento e, até mesmo, custos indiretos (luz, água, refrigeração, aluguel) em que se incorre para operar e abrigar recursos tecnológicos. O modelo de **custo total de propriedade (TCO – total cost of ownership)** pode ser utilizado para analisar esses custos diretos e indiretos, auxiliando as empresas a determinar o custo real de implantação de tecnologias específicas. A Tabela 4.6 descreve os componentes mais importantes do TCO a levar em conta quando se analisa esse custo.

Considerados todos os componentes, o TCO para um PC pode atingir até três vezes o preço original de compra do equipamento. Os ‘custos ocultos’ referentes a pessoal de suporte, *downtime* e gestão adicional da rede podem tornar as arquiteturas cliente/servidor distribuídas — especialmente as que incorporam computadores portáteis e dispositivos sem fio — mais caras do que as arquiteturas centralizadas com mainframes.

Muitas grandes empresas estão sobrecarregadas de recursos de hardware e software redundantes e incompatíveis, porque seus departamentos e divisões tiveram permissão para fazer suas próprias compras de tecnologia. Essas empresas poderiam reduzir seu TCO aumentando a centralização e a padronização dos recursos de hardware e software. Também poderiam reduzir o tamanho das equipes de sistemas de informação necessárias para apoiar essa infraestrutura, caso diminuíssem o número de modelos de computadores e produtos de software que os funcionários podem usar.

Utilização de provedores de serviços tecnológicos

Algumas das mais importantes perguntas que os administradores enfrentam são: “como devemos adquirir e manter nossos recursos tecnológicos? Devemos desenvolvê-los e executá-los por conta própria ou adquiri-los de fontes externas?”. No passado, a maioria das empresas desenvolvia e executava internamente as próprias instalações computacionais e desenvolvia o seu próprio software. Hoje, um número cada vez maior de empresas obtém sua tecnologia de hardware e software de provedores de serviços externos.

Tabela 4.6 Componentes do custo total de propriedade

Aquisição de hardware	Preço de compra do hardware do equipamento de computação, incluindo computadores, terminais, armazenamento, impressoras.
Aquisição de software	Compra ou licenciamento de software para cada usuário.
Instalação	Custos para instalar computadores e software.
Treinamento	Custos para dar treinamento a especialistas e usuários finais dos sistemas de informação.
Suporte	Custos para prover suporte técnico continuado, centrais de assistência e assim por diante.
Manutenção	Custos de atualização de hardware e software.
Infraestrutura	Custo para adquirir, manter e dar suporte à infraestrutura relacionada, tal como redes e equipamento especializado (inclusive unidades de cópia de segurança — backup).
<i>Downtime</i> (tempo em que o sistema fica indisponível, seja por manutenção ou por falhas)	Perda de produtividade caso falhas no hardware ou no software tornem o sistema indisponível para processamento e tarefas solicitadas pelos usuários.
Espaço e energia	Custos imobiliários e de instalações em que se incorre por abrigar a tecnologia e mantê-la operante.

Outsourcing

Inúmeras empresas estão delegando a fornecedores externos a manutenção de suas infraestruturas de TI e o desenvolvimento de novos sistemas, um processo denominado **outsourcing (terceirização)**. Elas podem fechar um contrato com um provedor de serviços externo para operar suas redes e seu centro computacional, desenvolver novos softwares ou gerenciar todos os componentes de sua infraestrutura de TI, como fez a Procter & Gamble (P&G). A P&G concordou em pagar à Hewlett-Packard (HP) 3 bilhões de dólares para gerenciar sua infraestrutura de TI, as operações de seu centro computacional, o suporte aos PCs e ao usuário final, a gestão de rede e o desenvolvimento e manutenção de aplicativos para suas operações globais em 160 países.

Empresas que não dispõem de recursos financeiros ou técnicos para operar seus próprios sites podem usar **serviços de Web hosting** especializados. Esses serviços mantêm um computador servidor Web de grande capacidade, ou baterias de servidores, e proveem espaço a assinantes para que possam manter seus sites mediante uma tarifa. As empresas inscritas podem criar suas páginas na Internet ou contratar o serviço de hosting ou de uma empresa especializada em design de página para criá-las. Alguns serviços oferecem **colocação**, na qual a empresa realmente é proprietária do computador servidor que hospeda seu site, mas o instala no espaço físico do serviço de hosting.

Muitas vezes, as empresas mantêm controle sobre seus recursos de hardware, mas terceirizam o desenvolvimento e a manutenção de software a outras empresas, com frequência a empresas que operam em países com baixos salários. Quando as empresas terceirizam atividades de software para fora das fronteiras nacionais, dizemos que elas estão fazendo um **outsourcing internacional de software (offshore software outsourcing)**. Até pouco tempo, esse tipo de desenvolvimento de software oferecia baixa qualidade em termos de manutenção, entrada de dados e operações de call center, mas, com a crescente sofisticação e experiência das empresas estrangeiras, particularmente as indianas, cada vez mais o desenvolvimento de novos programas ocorre em outros países. O Capítulo 11 discute o *outsourcing* internacional de software em mais detalhes.

Para gerenciar o relacionamento com as empresas estrangeiras, ou provedor de serviço de tecnologia, as empresas precisarão de um contrato que inclua um **acordo de nível de serviços (SLA – service level agreement)**, um contrato formal entre os clientes e seus provedores de serviços que define as responsabilidades específicas do provedor e o nível de serviço esperado pelo cliente. SLAs costumam especificar a natureza e o nível dos serviços fornecidos, critérios para avaliação de desempenho, opções de suporte, cláusulas para segurança e recuperação de desastres, propriedade e atualização de hardware e software, suporte ao cliente, cobrança, e condições para encerramento do acordo.

Utilizando os serviços da nuvem

Atualmente, as empresas têm a opção de manter sua própria infraestrutura de TI ou usar os serviços de hardware e software baseados na nuvem. Existem muitos fatores que devem ser considerados quando da decisão sobre manter uma infraestrutura de TI própria ou deixá-la a um fornecedor externo. As empresas que estiverem considerando o modelo da computação em nuvem precisam avaliar cuidadosamente os custos e os benefícios dos serviços externos, pesando todas as questões humanas, organizacionais e tecnológicas, incluindo o nível de serviço e desempenho aceitável para a empresa.

Empresas pequenas, que normalmente não dispõem dos recursos para manutenção de seus próprios hardware e software, podem achar muito mais fácil ‘alugar’ a infraestrutura de outra empresa e evitar os gastos e dificuldades com instalação, operação e manutenção de hardware e software próprios. Para as grandes empresas, os custos e as vantagens de negócios são menos claras. Elas terão de considerar não só os custos imediatos e de longo prazo de mudança para nuvem, mas também se os recursos são suficientemente seguros ou confiáveis para seus sistemas mais críticos. No momento, a maioria das grandes empresas é mais propensa a manter a maior parte de sua infraestrutura internamente, mas está disposta a mover gradualmente seu trabalho não essencial para os serviços da nuvem.

Um padrão emergente é descarregar a demanda de pico por serviços computacionais para grandes centros de dados remotos. Dessa forma, as empresas reduzem seus gastos com tecnologia e investem somente o necessário para gerenciar cargas médias de processamento e pagam só pelo poder computacional necessário conforme demandas do mercado. Essa organização permite que as empresas possuam infraestrutura mais flexível, com uma parte pertencente à empresa e outra alugada de centros computacionais gigantescos gerenciados por especialistas em tecnologia.

O software oferecido por provedores da nuvem pode ser adequado para funções básicas de desktop, mas não para a gestão de grandes sistemas corporativos com muitos processos de negócios exclusivos e complexos. Em alguns casos, o custo com o aluguel de software fica acima do custo da aquisição e manutenção internas da aplicação. Ainda assim, pode haver benefícios na utilização do software como serviço (SaaS) se ele permitir que a empresa se concentre nas questões de negócio essenciais em vez de pensar nos desafios tecnológicos.

Gerenciando plataformas móveis

Uma grande empresa pode ter milhares de dispositivos sem fio para configurar e monitorar, semelhante a um ambiente desktop. Será um desafio integrar essa nova plataforma à infraestrutura de TI e aplicações existentes na empresa. Coordenação e supervisão centrais são essenciais. As empresas precisarão preparar o inventário de todos os seus dispositivos móveis e desenvolver políticas e ferramentas para controle, atualização e segurança desses dispositivos, e para controle dos dados e aplicações que neles funcionam.

Ganhos na produtividade e na eficiência trazidos pela distribuição de dispositivos de computação móveis aos empregados devem ser contrabalançados com o aumento nos custos com integração desses dispositivos à infraestrutura de TI da empresa e com suporte técnico. Outros componentes de custos incluem taxas para uso de conexão sem fio, treinamento do usuário final, suporte e software para aplicações específicas.

Embora os custos de um PDA sem fio para um empregado corporativo possam chegar a algumas centenas de dólares, o TCO para cada dispositivo é muito mais alto, variando de 1 a 3 mil dólares, conforme estimativas de vários especialistas. Os custos são mais altos se os dispositivos móveis rodam muitas aplicações diferentes ou precisam ser integrados a sistemas de *back-end*, como aplicações corporativas.

Gerenciando a localização de software para empresas globais

Se você está gerenciando uma empresa global, todas as questões gerenciais já descritas serão afetadas pela necessidade de criação de sistemas que possam ser realisticamente utilizados por múltiplas unidades de negócios em diferentes países. Embora o inglês tenha se tornado uma espécie de língua-padrão aos negócios, essa afirmação é mais verdadeira nos níveis mais altos da empresa do que nas camadas média e baixa. Pode ser que o software precise ser construído com interfaces no idioma local antes que um novo sistema de informação possa ser implantado em todo o mundo.

Pode ser caro e confuso construir essas interfaces. Menus, comandos, mensagens de erro, relatórios, consultas, formulários para entrada de dados on-line e documentação do sistema podem precisar ser traduzidos para todas as línguas dos países nos quais o sistema será utilizado. Para ser realmente útil no aumento da produtividade de uma força de trabalho global, as interfaces do software devem ser facilmente compreendidas e aprendidas rapidamente. Todo o processo de conversão de software para funcionamento em uma segunda língua é denominado *localização de software*.

Os sistemas globais devem ainda considerar as diferenças nas culturas locais e nos processos de negócios. Sistemas multifuncionais, como os de gestão empresarial e os sistemas de gestão da cadeia de abastecimento, nem sempre são compatíveis com as diferenças em idiomas, heranças culturais e processos de negócios em outros países.

Em um ambiente de sistemas globais, todos esses fatores aumentam o TOC e influenciam nas decisões sobre terceirizar ou utilizar os provedores de serviços de tecnologia.

Projetos práticos em SIG

Os projetos nesta seção oferecem experiência prática na identificação de sistemas de informação para suporte à estratégia de negócios e para resolver um problema de retenção de clientes usando banco de dados para melhorar a tomada de decisão em relação à estratégia de negócios, e utilizando ferramentas da Web para projetar e cotar um automóvel.

Problemas de decisões gerenciais

1. O Centro Médico da Universidade de Pittsburgh conta com sistemas de informações para operar 19 hospitais, uma rede de outros locais de cuidados com a saúde e empreendimentos comerciais e internacionais. A demanda por novos servidores e tecnologia de armazenamento estava crescendo 20 por cento a cada ano. O Centro configurava um servidor separado para cada aplicação, e seus servidores e outros computadores estavam funcionando sob diferentes sistemas operacionais, inclusive diversas versões de UNIX e Windows. Era preciso gerenciar tecnologias de muitos fornecedores diferentes, como Hewlett-Packard (HP), Sun Microsystems, Microsoft e IBM. Avalie o impacto dessa situação sobre o desempenho dos negócios. Quais fatores e decisões gerenciais devem ser considerados quando do desenvolvimento de uma solução para o problema?
2. A Qantas Airways, empresa aérea líder na Austrália, enfrenta pressões nos custos com os altos preços do combustível e os baixos níveis de tráfego aéreo global. Para continuar competitiva, a empresa deve encontrar maneiras de manter os custos baixos enquanto oferece alto nível de atendimento ao cliente. A Qantas possuía centro de dados há 30 anos. A gerência precisava decidir entre a atualização de sua infraestrutura de TI com novas tecnologias ou a terceirização. Que fatores devem ser considerados pela gerência da corporação se optar pela terceirização? Liste e descreva os pontos a serem considerados no acordo de nível de serviço se a organização optar pela terceirização.

Aperfeiçoando a tomada de decisões: use uma planilha eletrônica para avaliar opções de hardware e software

Habilidades de software: fórmulas em planilha eletrônica

Habilidades organizacionais: pesquisa de preço de tecnologia

Neste exercício, você vai usar software de planilha eletrônica para calcular o custo de sistemas desktop alternativos.

Solicitaram que pesquisasse o preço de equipamentos de hardware e software para um escritório com 30 pessoas. Usando a Internet, pesquise o preço de 30 PCs (incluindo monitor, computador e teclado) da Lenovo, da Dell e da HP/Compaq, conforme anunciados no site corporativo de cada fabricante. (Para os propósitos deste exercício, ignore o fato de que os PCs normalmente vêm com pacotes de software pré-instalados.) Pesquise também o preço de 15 impressoras a laser monocromáticas para PC da HP e da Xerox. Cada PC precisa satisfazer as especificações mínimas apresentadas na tabela a seguir:

Especificações mínimas para PC	
Velocidade do processador	3GHz
Disco rígido	250 GB
RAM	3 GB
Velocidade do DVD-ROM	16X
Monitor (medida diagonal)	17 polegadas

Cada impressora precisa satisfazer as especificações mínimas apresentadas na tabela a seguir:

Especificações mínimas para impressora a laser

Velocidade de impressão	12 páginas por minuto
Resolução de impressão	600 x 600
Pronta para trabalhar em rede?	Sim
Preço máximo/unidade	\$ 700,00

Depois de pesquisar o preço dos PCs e das impressoras, pesquise o preço de 30 cópias das versões mais recentes do Microsoft Office e do Sun StarOffice, e de 30 cópias do Microsoft Windows 7 Professional. Os pacotes de software aplicativos possuem várias versões, portanto, certifique-se de que cada pacote contenha programas para processamento de texto, análise de planilha eletrônica, análise de banco de dados, recursos de apresentação gráfica e e-mail.

Prepare uma planilha eletrônica apresentando os resultados de sua pesquisa quanto aos PCs, impressoras e softwares. Use seu software de planilha para determinar qual combinação de PC, impressora e software oferece, ao mesmo tempo, melhor desempenho e preço mais baixo por funcionário. Como cada dupla de funcionários dividirá a mesma impressora (15 impressoras/30 PCs), coloque na planilha apenas a metade do custo da impressora por funcionário. Suponha que sua empresa usará a garantia e o contrato de serviços-padrão oferecidos pelo fabricante.

Aperfeiçoando a tomada de decisões: use a pesquisa na Web para planejamento e orçamento de uma conferência de vendas

Habilidades de software: software baseado na Internet

Habilidades organizacionais: pesquisa de preços de transportes e hospedagem

Neste exercício, você irá utilizar programas em diferentes sites de viagem para conseguir transporte e hospedagem para uma grande força de vendas que irá a uma conferência de vendas em dois locais possíveis. Você usará a informação obtida para calcular os custos totais de viagem e hospedagem e decidir onde realizar a conferência.

A Foremost Composite Materials Company está planejando realizar uma conferência de vendas entre 19 e 20 de outubro, que começará com uma recepção na noite de 18 de outubro. A conferência consiste em encontros que duram o dia todo e aos quais toda a força de vendas — 125 representantes de vendas e seus 16 gerentes — precisa comparecer. Cada representante de vendas ocupará um quarto *single*; serão necessárias duas salas de reunião, uma grande o suficiente para abrigar toda a força de vendas mais alguns visitantes (200 pessoas), e a outra para a metade da força. A administração designou uma verba de 110 mil dólares para as diárias dos representantes. O hotel também precisa oferecer diversos serviços como retroprojeter e *data show*, assim como restaurante e centro para negócios. Além desses serviços, também oferecerá infraestrutura para que possam trabalhar no quarto e desfrutar de piscina ou academia nas horas vagas. A empresa gostaria de realizar a conferência em Miami ou em Marco Island, na Flórida.

A Foremost normalmente realiza congressos desse gênero nos hotéis Hilton ou Marriott. Use os sites dessas redes para escolher um hotel, em qualquer uma das cidades mencionadas, que permita à empresa realizar a conferência dentro de seu orçamento.

Acesse a home page de cada rede e pesquise-a para encontrar um hotel que atenda aos requisitos da Foremost. Selecionado o hotel, localize os voos com chegada na noite anterior à conferência; afinal, os participantes precisam fazer o check-in e comparecer à recepção antes da conferência. Os participantes virão de Los Angeles (54), San Francisco (32), Seattle (22), Chicago (19) e Pittsburgh (14). Determine quais serão os custos de cada bilhete aéreo a partir dessas cidades. Ao terminar seu levantamento, crie um orçamento para a conferência. O orçamento deve incluir o custo do bilhete aéreo, da diária do hotel e os 60 dólares diários por participante para alimentação.

- Qual seu orçamento final?
- Que hotel escolheria para a conferência de vendas e por quê?

Resumo

1 Identifique e descreva os componentes da infraestrutura de TI. Por infraestrutura de TI, entendem-se os recursos tecnológicos compartilhados que formam a plataforma sobre a qual os aplicativos de sistema de informação específicos da empresa se apoiam. A infraestrutura de TI inclui hardware, software e serviços compartilhados por toda a empresa. Entre os principais componentes da infraestrutura de TI estão hardware, software, tecnologia de armazenamento de dados, tecnologia de rede e telecomunicações e serviços de tecnologia.

2 Identifique e descreva os principais tipos de tecnologia de hardware e de armazenamento, entrada e saída de dados. Os computadores podem ser classificados em mainframes, computadores de médio porte, PCs, estações de trabalho e supercomputadores. Os mainframes são os computadores de maior porte; os computadores de médio porte são servidores; PCs são computadores de mesa ou laptops; estações de trabalho são computadores de mesa com potentes recursos gráficos e matemáticos; e supercomputadores são potentes e sofisticados computadores que podem realizar rapidamente operações complexas e numerosas. A capacidade computacional pode ainda ser aumentada caso computadores geograficamente distantes sejam conectados em uma única rede, criando-se assim uma computação em grade que agrega a capacidade de todos os computadores da rede. No modelo de computação cliente/servidor, o processamento é dividido entre 'clientes' e 'servidores' conectados por uma rede. A divisão exata de tarefas entre cliente e servidor depende do aplicativo.

As principais tecnologias de armazenamento secundário são os discos magnéticos, os discos ópticos e as fitas magnéticas. Os discos ópticos (CD-ROMs e DVDs) podem armazenar grandes quantidades de dados de maneira compacta, sendo alguns tipos regraváveis. As redes de armazenamento de dados (SANs) conectam múltiplos dispositivos de armazenamento em uma rede separada de alta velocidade dedicada apenas à armazenagem. Os principais dispositivos de entrada são teclado, mouse de computador, tela sensível ao toque, reconhecimento óptico de caracteres, reconhecimento de caracteres de tinta magnética (MICR), entrada por caneta, dispositivo de varredura digital (scanner digital), sensores, dispositivos de entrada de áudio e dispositivos de identificação por radiofrequência. Os principais dispositivos de saída são monitores, impressoras e dispositivos de saída de áudio.

3 Identifique e descreva os principais tipos de software usados pelas empresas. Existem dois tipos principais de software: o software de sistema e o software aplicativo. O software de sistema coordena as várias partes do sistema computacional e faz a intermediação entre o software aplicativo e o hardware. O software aplicativo é usado para desenvolver aplicativos empresariais específicos.

O software de sistema que gerencia e controla as atividades do computador é denominado sistema operacional. Entre os principais sistemas operacionais para PC e servidores estão Windows Vista, Windows 7, Windows Server 2008, UNIX, Linux e o sistema operacional para Macintosh. O Linux é um sistema operacional poderoso, resiliente e de código aberto que pode rodar em inúmeras plataformas de hardware; é amplamente utilizado em servidores Web.

As principais linguagens de programação usadas em softwares aplicativos empresariais são COBOL, C, C++ e Visual Basic. Menos orientadas a procedimentos do que as linguagens convencionais, as de quarta geração permitem aos usuários finais realizar muitas tarefas de software que anteriormente exigiam técnicos especializados. Entre elas estão as populares ferramentas de software para PC e para a nuvem, tais como editores de texto, planilhas eletrônicas, gestão de dados, apresentações gráficas e softwares de e-mail, além de navegadores da Web. Java, uma linguagem de programação não vinculada a um hardware ou sistema operacional específico, é hoje o principal ambiente de programação interativo para a Web. HTML é uma linguagem de descrição de página para a criação de páginas da Web.

Os serviços Web são componentes de software vagamente relacionados, baseados em XML e padrões da Web abertos, não vinculados a um produto específico e podem trabalhar com qualquer software aplicativo e sistema operacional. Podem ser usados como componentes de aplicativos da Internet que integrem os sistemas de duas organizações distintas ou os sistemas dispares de uma mesma empresa.

4 Enumere as tendências atuais de hardware e software. As tendências atuais de hardware e software demonstram que, cada vez mais, a computação está acontecendo em plataformas móveis digitais. A computação em nuvem oferece recursos de hardware e software na forma de serviços disponibilizados pela Internet. Na computação autônoma, sistemas computacionais contam com recursos para automaticamente configurar e consertar a si mesmos. O software de código aberto está se proliferando porque permite que os usuários o modifiquem como quiserem e utilizem-no como plataforma para novas aplicações derivadas. *Mashups* e *widgets* são blocos de construção utilizados na criação de novos aplicativos e serviços de software utilizando o modelo da computação em nuvem. O software como serviço (SaaS) distribui software remotamente como um serviço sob demanda da rede.

5 Avalie as principais questões envolvidas na administração dos recursos de hardware e software. Administradores e especialistas em sistemas de informação precisam prestar especial atenção ao planejamento da capacidade e da escalabilidade do hardware, asse-

gurando que a empresa tenha capacidade computacional suficiente para suas necessidades atuais e futuras. As empresas também precisam pesar os custos e benefícios de desenvolver e manter seu próprio hardware e software ou, alternativamente, terceirizar ou utilizar o modelo de computação sob demanda. O custo total de propriedade

(TCO) dos recursos tecnológicos inclui não apenas o custo original do hardware e do software, mas também os relacionados a atualizações, manutenção, suporte técnico e treinamento, inclusive os custos para gestão e manutenção de dispositivos móveis. Empresas com operações globais precisam ainda gerenciar a localização de software.

Palavras-chave

Acordo de nível de serviços (SLA), 132	HTML (<i>Hypertext Markup Language</i>), 126	Redes de armazenamento de dados (SANs), 109
Ajax, 125	Interface gráfica de usuário (GUI), 117	Serviços Web, 126
Arquiteturas cliente/servidor multicamadas (N-camadas), 108	Java, 125	Servidor, 106
Arquitetura orientada a serviços (SOA), 127	Linguagem de programação visual, 121	Servidor Web, 108
C, 121	Linguagens de consulta, 121	Servidor de aplicativo, 108
C++, 121	Linguagens de quarta geração, 121	Sistema operacional, 117
CD-ROM (disco compacto com memória somente de leitura), 109	Linux, 120	Sistemas legados, 106
Cliente, 107	Mainframe, 106	Software como serviço (SaaS), 129
Computação autônoma, 114	<i>Mashups</i> , 129	Software de código aberto, 120
Computação cliente/servidor, 107	Multitoque, 117	Software de gestão de dados, 105
Computação em grade, 107	Nanotecnologia, 110	Software de planilha eletrônica, 123
Computação em nuvem, 112	Navegadores da Web, 125	Software de processamento de textos, 123
Computação sob demanda, 113	Netbooks, 110	Softwares aplicativos, 105
Computador pessoal (PC), 106	Objeto, 121	Softwares de sistema, 105
Custo total de propriedade (TCO), 131	Office 2010, 124	Supercomputador, 107
Disco magnético, 109	<i>Outsourcing</i> , 132	UNIX, 120
Discos de vídeo digital (DVDs), 109	<i>Outsourcing</i> internacional de software, 132	Virtualização, 114
Dispositivos de entrada, 109	Pacote de software, 121	Visual Basic, 121
Dispositivos de saída, 109	Planejamento da capacidade, 130	Widget, 129
Escalabilidade, 130	Processador multinúcleo, 116	Windows 7, 119
Estação de trabalho, 106	Processamento centralizado, 107	Windows Vista, 119
Fita magnética, 109	Processamento distribuído, 107	XML (<i>eXtensible Markup Language</i>), 126
	Software de apresentação gráfica, 123	

Questões de revisão

- Quais os componentes da infraestrutura de tecnologia de informação (TI)?
 - Defina infraestrutura de tecnologia de informação (TI) e descreva cada um de seus componentes.
- Quais os principais dispositivos de hardware, armazenamento de dados, entrada e saída utilizados nas empresas?
 - Enumere e descreva os vários tipos de computadores disponíveis para as empresas atualmente.
 - Defina o modelo de computação cliente/servidor e descreva a diferença entre uma arquitetura cliente/servidor de duas camadas e de N-camadas.
 - Enumere os mais importantes meios de armazenagem secundária e descreva as vantagens e as limitações de cada um.
 - Enumere e descreva os principais dispositivos de entrada e saída computacional.
- Quais os principais tipos de software utilizados nas empresas?
 - Descreva a diferença entre software aplicativo e de sistema e explique o papel desempenhado pelo sistema operacional de um computador.
 - Enumere e descreva os principais sistemas operacionais para PCs e para servidores.
 - Nomeie e descreva cada tipo de ferramenta de software de quarta geração e explique no que as linguagens de quarta geração diferem das linguagens de programação convencionais.
 - Nomeie e descreva as principais ferramentas de software de produtividade para PCs.
 - Explique como as linguagens Java e HTML são usadas no desenvolvimento de aplicativos para a Web.
 - Defina o que são serviços da Web, descreva a tecnologia que utilizam e explique como podem ajudar as empresas.

4. Quais são as tendências de hardware e software mais importantes?

- Defina e descreva plataforma digital móvel, nanotecnologia, computação em grade, computação em nuvem, computação autônoma, virtualização e processamento multinúcleo.
- Explique por que o software de código aberto é tão importante atualmente e seus benefícios para os negócios.
- Enumere e descreva serviços de software da computação em nuvem, *mashups* e *widgets*, e explique como indivíduos e empresas podem se beneficiar dos mesmos.

5. Quais as principais preocupações na gestão da tecnologia de hardware e software?

- Explique por que os gerentes precisam estar atentos ao planejamento da capacidade e à escalabilidade dos recursos tecnológicos.
- Descreva os componentes de custo usados para calcular o TCO dos ativos tecnológicos.
- Identifique os benefícios e desafios da utilização de *outsourcing*, serviços de computação em nuvem e plataformas móveis.
- Explique por que a localização de software se tornou uma preocupação importante de gestão para as empresas globais.

Para discutir

1. Por que a escolha do hardware e do software para a organização é uma importante decisão empresarial? Quais questões humanas, organizacionais e tecnológicas precisam ser consideradas durante essa escolha?

2. As organizações devem usar provedores de serviços aplicativos, inclusive serviços de computação em nuvem, para todas as suas necessidades de software? Por quê? Quais fatores humanos, organizacionais e tecnológicos devem ser considerados quando se toma essa decisão?

Colaboração e trabalho em equipe

Avaliando sistemas operacionais móveis e para servidores

Forme um grupo de três ou quatro colegas. Escolha e avalie sistemas operacionais móveis e para servidores. Vocês podem comparar os recursos e custos do Linux em relação à mais recente versão do sistema operacional Windows para servidores ou em relação ao UNIX. Vocês também podem comparar o sistema operacional móvel Android com o Symbian, ou qualquer um desses dois

com a versão mais recente do sistema operacional para iPhone (iOS). Se possível, use o Google Sites para postar links para outras páginas da Web, anúncios para a equipe, trabalhos; para trocar ideias e trabalhar colaborativamente em documentos do projeto. Tente usar o Google Docs para desenvolver uma apresentação de suas descobertas para sua turma.

Resolvendo problemas organizacionais

Salesforce.com: serviços de software em nuvem viram uma tendência

A Salesforce.com Inc., uma das maiores empresas de tecnologia disruptiva das últimas décadas, estremeceu a indústria de software com seu modelo de negócios inovador e de grande sucesso. A Salesforce oferece gestão do relacionamento com clientes (CRM) e outras soluções de software aplicativo na forma de 'software como serviço' (SaaS) distribuído pela Internet, em oposição ao software comprado e instalado localmente em computadores.

A empresa foi fundada em 1999 por Marc Benioff, ex-executivo da Oracle, conta com 2.600 empregados e, em 2009, gerou 1,077 bilhão de dólares em receita, fazendo com que ela se tornasse uma das 50 maiores empresas de software do mundo. A Salesforce.com possui mais de 55 mil clientes corporativos e mais de 1,5 milhão de assinantes. A empresa atribui seu sucesso aos muitos benefícios de seu modelo de distribuição de software sob demanda.

O modelo sob demanda elimina as necessidades de altos investimentos iniciais em sistemas e longas implan-

tações de computadores corporativos. O valor das assinaturas começa em 9 dólares por usuário ao mês pela versão de grupo reduzida para pequenas equipes de vendas e marketing, e a assinatura mensal mais barata para versões mais avançadas destinadas às grandes empresas custa 65 dólares por usuário.

A Häagen-Dazs Shoppe de Minneapolis, pertencente à Nestlé USA, por exemplo, calculou que teria de gastar 65 mil dólares com um banco de dados customizado para ajudar a gerência a manter contato com as franquias de varejo da empresa. Para usar os serviços da Salesforce, a empresa precisou pagar somente 20 mil dólares, mais uma taxa mensal de 125 dólares para que 20 usuários utilizassem PDAs sem fio, ou a Internet, para monitorar todas as franquias Häagen-Dazs nos Estados Unidos.

As implantações da Salesforce.com levam, no máximo, três meses, mas normalmente são concluídas em menos de um mês. Os compradores não precisam adquirir, esca-

lar ou manter nenhum tipo de hardware. Não é necessária a instalação de sistemas operacionais, bancos de dados ou servidores de aplicação; e não são exigidos consultores, equipes, licenças caras ou taxas de manutenção. O sistema pode ser acessado através de um navegador-padrão da Web e a Salesforce.com atualiza constantemente o software nos bastidores. Existem ferramentas para customização de alguns recursos do software para suportar os processos de negócios exclusivos de determinadas empresas. Os assinantes não precisam realizar investimentos iniciais massivos em hardware e software, e podem cancelar a assinatura se os negócios forem mal ou se surgir um sistema melhor. Se dispensarem empregados, as empresas podem ainda diminuir o número de assinaturas contratadas.

A Salesforce enfrenta desafios importantes à medida que continua a crescer e aprimorar seus negócios. O primeiro refere-se ao aumento na concorrência tanto em relação aos líderes tradicionais do setor quanto às novas empresas que sonham em reproduzir o sucesso da Salesforce. Microsoft, SAP e Oracle lançaram versões baseadas em assinaturas de seus produtos CRM. Concorrentes menores, como NetSuite, Salesboom.com e RightNow, também já começaram a invadir o espaço de mercado da Salesforce.

A Salesforce ainda tem muito a fazer para alcançar o tamanho e o mercado de seus maiores concorrentes. Em 2007, a fatia de mercado da SAP era quase quatro vezes maior do que a da Salesforce.com, e a base de clientes da IBM conta com 9 mil empresas de software que rodam suas aplicações em produtos IBM e estão pouco propensas a escolher a Salesforce como substituta.

É preciso que a Salesforce continue a provar a seus clientes que é confiável e segura o suficiente para gerenciar remotamente seus dados e aplicações corporativas. A organização enfrentou uma série de panes, sendo a mais recente delas no início de 2009. Em 6 de janeiro daquele ano, um dispositivo de rede essencial falhou e, por 38 minutos, impediu que os dados oriundos da Europa, do Japão e dos Estados Unidos fossem processados. Mais de 177 milhões de transações foram afetadas. Embora a maioria dos clientes da Salesforce compreenda que os serviços de TI disponibilizados pela nuvem não estarão disponíveis em tempo integral, alguns clientes e críticos aproveitaram as falhas para questionar a estabilidade de todo o conceito de computação em nuvem. Uma falha semelhante ocorreu em fevereiro, afetando a Europa e, algumas horas depois, os Estados Unidos.

Até agora, a Salesforce.com vivenciou somente uma violação à segurança. Em novembro de 2007, um empregado da empresa foi vítima de um ataque de *phishing* e divulgou sua senha corporativa a fraudadores, o que expôs a lista de clientes da Salesforce. Os clientes da empresa foram alvo de uma enxurrada de golpes altamente direcionados e de tentativas ilegais de acesso que pareceram autênticas. Embora esse incidente tenha levantado uma bandeira vermelha, muitos clientes relataram que a solução da Salesforce foi satisfatória. Os principais clientes normalmente enviam auditores para verificar a segurança da organização.

Outro desafio é expandir seu modelo de negócios para outras áreas. Os serviços costumam ser usados, em sua

maioria, por equipes de vendas que precisam controlar as perspectivas de vendas e a lista de clientes. A empresa está tentando oferecer funcionalidades adicionais através de uma parceria com a Google, mais especificamente como o serviço Google Apps. A Salesforce.com está combinando seus serviços com Gmail, Google Docs, Google Talk e Google Calendar para permitir que seus clientes realizem mais tarefas pela Web. A empresa também lançou uma ferramenta de desenvolvimento para integração ao Facebook. Os clientes podem construir aplicações que chamam funções do site Facebook.

A corporação também estabeleceu parceria com a Apple para distribuir suas aplicações ao uso no iPhone. A empresa espera conseguir entrar no amplo mercado de usuários do iPhone, lançando a possibilidade de uso das aplicações da Salesforce no celular a qualquer hora e lugar.

A parceria entre Salesforce e Google representa uma frente unida contra a Microsoft, cuja intenção é diminuir a popularidade do Microsoft Office. Atualmente, a Salesforce.com descreve a relação como "primeiramente um acordo de distribuição", mas que pode se fortalecer com base na ideia de que as empresas preferem gerenciar o relacionamento com os clientes e atividades relacionadas em um só lugar. Google e Salesforce.com esperam que a iniciativa de lançar uma Salesforce.com para o Google Apps estimule o futuro crescimento na área de software sob demanda.

Em fevereiro de 2009, rumores espalhavam que a Oracle estava disposta a comprar a Salesforce. Estudos especulavam que a aquisição seria boa para a Oracle porque a empresa ainda levaria muito tempo para desenvolver suas ofertas de software como serviço e chegar ao ponto em que a Salesforce já havia chegado. Até agora, no entanto, o negócio não se materializou, mas alguns observadores do setor esperam que, em algum momento, a Oracle acabe comprando a Salesforce.

Para aumentar sua receita e chegar aos níveis que os especialistas e Wall Street esperam, a empresa precisa tirar seu foco da venda de uma suíte de aplicações de software e mudá-lo para a oferta de uma plataforma de computação em nuvem mais ampla, na qual diversas empresas de software possam oferecer suas aplicações. Segundo o CEO Marc Benioff, ao longo da última década, "nos concentramos em software como serviço". Na próxima década, a Salesforce.com "irá realmente se concentrar na plataforma como serviço".

A empresa intensificou seus esforços para fornecer ofertas de computação em nuvem a seus clientes. O novo site enfatiza muito mais a computação em nuvem, agrupando os produtos em três tipos: nuvem de vendas, de serviços e customizada. As nuvens de vendas e de serviços consistem de aplicações com a intenção de melhorar as vendas e o atendimento aos clientes, respectivamente, e a nuvem customizada é o outro nome para a plataforma de desenvolvimento Force.com, na qual os clientes podem desenvolver suas próprias aplicações para uso na ampla rede Salesforce.

A plataforma Force.com oferece um conjunto de ferramentas de desenvolvimento e serviços de TI que permitem ao usuário customizar suas aplicações Salesforce. com de gestão do relacionamento com os clientes ou

construam aplicações completamente novas para serem executadas “na nuvem” da infraestrutura do centro de dados da Salesforce.com. A empresa abriu o Force.com para outros desenvolvedores independentes de software e listou seus programas em seu AppExchange.

Utilizando AppExchange, as pequenas organizações podem se conectar à Internet e baixar facilmente mais de 800 aplicações de software, alguns complementos aos serviços Salesforce.com e outros não relacionados, mesmo em funções não direcionadas ao relacionamento com o cliente, como recursos humanos. Os sites Force.com, baseados no ambiente de desenvolvimento Force.com, permitem que os usuários desenvolvam páginas Web e registrem nomes de domínios. O preço baseia-se no tráfego do site.

A infraestrutura de nuvem da Salesforce.com inclui dois centros de dados nos Estados Unidos e um terceiro em Cingapura, e há planejamento de outros na Europa e no Japão. A Salesforce.com também estabeleceu parceria com a Amazon para permitir que os clientes da plataforma Force.com utilizassem os serviços de computação em nuvem da Amazon (Elastic Compute Cloud e Simple Storage Service). Os serviços da Amazon gerenciam as tarefas das aplicações Force.com que demandam poder computacional ou capacidade de armazenamento adicionais.

A Author Solutions de Bloomington, Minnesota, usa a plataforma Force.com para hospedar as aplicações por trás de suas operações e expandir sua biblioteca de ideias. A Authors.com é a maior empresa de autopublicação no mundo e lançou 24 mil títulos em 2008, com receita anual de 100 milhões de dólares e 400 empregados. O negócio está crescendo a uma taxa acelerada e já comprou diversas outras empresas. A Author Solutions esforçava-se para gerenciar dois bancos de dados, dois sistemas de comércio eletrônico e três sistemas de fluxo de trabalho para apoiar suas três marcas. A gerência sênior decidiu consolidar tudo isso em uma única plataforma.

A Salesforce.com foi a plataforma escolhida. A Author Solutions admirou a possibilidade de escalar à medida que a empresa crescia e a economia de 75 por cento, pois não teve de manter e gerenciar seu próprio centro de dados. As modificações no fluxo de trabalho, que já levaram de 30 a 120 horas para serem concluídas, são realizadas em um quarto do tempo. O tempo e o custo para inclusão de um novo produto, que costumava levar de 120 a 240 horas e custava entre 6 mil e 12 mil dólares, foram reduzidos em 75 por cento. Cinco meses depois de

assinar com a Salesforce, a Author Solutions já estava utilizando seu novo “sistema de planejamento dos recursos de publicação”, que coordena todos os processos envolvidos na contratação e publicação de um livro — controle de projeções, gestão de submissões editoriais e correções. A nova plataforma consegue gerenciar um volume de trabalho 30 por cento maior que o antigo sistema, com o mesmo número de empregados.

A questão é se o público para as plataformas AppExchange e Force.com, da Salesforce.com, será grande o suficiente para acarretar o nível de crescimento esperado pela empresa. Ainda não está claro se irá gerar a receita da qual precisa para oferecer serviços de computação em nuvem na mesma escala que o Google ou a Amazon, e ainda recuperar seus investimentos no novo modelo.

Alguns analistas acreditam que a plataforma pode não ser atrativa para as grandes empresas e suas necessidades de aplicação. Outro desafio é oferecer disponibilidade constante. Os assinantes da Salesforce.com precisam que o serviço esteja disponível em tempo integral. Graças às falhas relatadas anteriormente, entretanto, muitas organizações repensaram sua dependência do software como serviço. A Salesforce.com oferece ferramentas para garantir aos clientes a confiabilidade de seu sistema e também disponibiliza aplicações para PC que se relacionam a seus serviços de forma que os usuários possam trabalhar off-line.

Ainda assim, um grupo de empresas continua relutante com relação ao software como serviço e à popularidade da computação em nuvem. Além disso, ainda não está claro se o software distribuído pela Internet custará menos no longo prazo. Segundo Robert DiSisto, analista da Gartner, pode ser mais barato ser um assinante dos serviços da Salesforce.com nos primeiros anos, mas o que acontece depois disso? Os custos de atualização e gestão do software sob demanda irão ultrapassar as economias com a utilização dos serviços da Salesforce.com?

Fontes: Clint Boulton, “Salesforce.com Network Device Failure Shuts Thousands Out of SaaS Apps”. *eWeek*, 7 jan. 2009; Jessi Hempel, “Salesforce Hits Stride”. *CNN Money.com*, 2 mar. 2009; Charles Babcock, “Salesforce Keeps Tight Control on Cloud Development”. *Information Week*, 27 mar. 2009; “More Books Fewer Worries”. *Customer Relationship Management*, ago. 2009; J. Nicholas Hoover, “Service Outages Force Cloud Adopters to Rethink Tactics”. *Information Week*, 18/25 ago. 2008; Saul Hansell, “How Apple and Facebook Influence Salesforce.com”. *The New York Times*, 17 dez. 2008; Charles Babcock, “Salesforce Ascends Beyond Software As Service”. *Information Week*, 10 nov. 2008; Mary Hayes Wejer, “Salesforce, Google Show Fruits of Their Collaboration”. *Information Week*, 21 abr. 2008; John Pallatto; Clint Boulton, “An On-Demand Partnership”. *eWeek*, 21 abr. 2008; Gary Rivlin, “Software for Rent”. *The New York Times*, 13 nov. 2007.

Questões de estudo de caso

1. Quais as vantagens e desvantagens do modelo de computação em nuvem?
2. Quais serão alguns dos desafios enfrentados pela Salesforce à medida que continuar crescendo? Quão bem a empresa conseguirá vencer esses desafios?
3. Que tipos de negócios poderiam se beneficiar com a mudança para a Salesforce? Explique.
4. Que fatores você consideraria se optasse por utilizar os serviços da Salesforce em sua empresa?
5. Uma empresa poderia administrar todo o seu negócio utilizando Salesforce.com, Force.com e App Exchange? Explique.

Referências bibliográficas

- BABCOCK, Charles. "Linux No Longer the Cool New Kid on the Block. Now What?". *Information Week*, 14 abr. 2008.
- BARRY, Douglas K. *Web Services and Service-Oriented Architectures: The Savvy Manager's Guide*. New York: Morgan Kaufman, 2003.
- CARR, David F. "Scaling Up or Scaling Out?". *Baseline*, jan. 2008.
- CARR, Nicholas. *The Big Switch*. New York: Norton, 2008.
- CHCKOWSKI, Ericka. "How Good Are Your Service-Level Agreements?". *Baseline*, jan. 2008.
- COLE, Arthur. "Mainframes They Are A'Changin". *ITBusinessEdge.com*, acesso em 20 jan. 2009.
- CONE, Edward. "The Grid Wins". *CIO Insight*, jan. 2008.
- DAVID, Julie Smith; SCHUFF, David; ST. LOUIS, Robert. "Managing Your IT Total Cost of Ownership". *Communications of the ACM*, v. 45, n. 1, jan. 2002.
- DUBEY, Abhijit; WAGLE, Dilip. "Delivering Software as a Service". *McKinsey Quarterly*, jun. 2007.
- ERICKSON, Jonathan. "Dr. Dobb's Report: Mobile Platforms". *Information Week*, 30 mar. 2009.
- FITZGERALD, Brian. "The Transformation of Open Source Software". *MIS Quarterly*, v. 30, n. 3, set. 2006.
- FOLEY, John. "IBM Cloud Shapes Up". *Information Week*, 22 jun. 2009.
- FOX, Armando; PATTERSON, David. "Self-Repairing Computers". *Scientific American*, maio 2003.
- GANEK, A. G.; CORBI, T. A. "The Dawning of the Autonomic Computing Era". *IBM Systems Journal*, v. 42, n. 1, 2003.
- HAGEL III, John; BROWN, John Seeley. "Your Next IT Strategy". *Harvard Business Review*, out. 2001.
- HELFT, Miguel; VANCE, Ashlee. "Google Plans a PC Operating System". *The New York Times*, 8 jul. 2009.
- HELFT, Miguel. "Reaching for the Cloud". *The New York Times*, 12 jul. 2009.
- HOOVER, J. Nicholas; MARTIN, Richard. "Demystifying the Cloud". *Information Week*, 23 jun. 2008.
- IBM. "Gridlines: The Intersection of Technology and Business". <http://www-1.ibm.com/grid/gridlines/January2004/feature/team-work.shtml>, acesso em jul. 2004.
- IBM. "IBM Launches New Autonomic Offerings for Self-Managing IT Systems". *IBM Media Relations*, 30 jun. 2005.
- KING, Rachael. "How Cloud Computing is Changing the World". *Business Week*, 4 ago. 2008.
- KONTZER, Tony. "The Cloud's the Limit: Anything as A Service". *Baseline*, jun. 2009.
- LAWTON, Christopher; CLARK, Don. "'Virtualization' is Pumping Up Servers". *The Wall Street Journal*, 6 mar. 2007.
- LOHR, Steve. "I.B.M. to Help Clients Fight Cost and Complexity". *The New York Times*, 15 jun. 2009.
- MARKOFF, John. "After the Transistor, a Leap into the Microcosm". *The New York Times*, 31 ago. 2009.
- McAFEE, Andrew. "Will Web Services Really Transform Collaboration?". *MIT Sloan Management Review*, v. 46, n. 2, inverno 2005.
- MEARIAN, Lucas. "A Zettabyte by 2010: Corporate Data Grows Fiftyfold in Three Years". *Computerworld*, 6 mar. 2007.
- MILLARD, Elizabeth. "The State of Mobile Applications". *Baseline*, 20 ago. 2008.
- PATEL, Samir; SAIGAL, Suneel. "When Computers Learn to Talk: A Web Services Primer". *McKinsey Quarterly*, n. 1, 2002.
- RAPOZA, Jim. "Browsers Battle for Enterprise Dominance". *eWeek*, 15 jun. 2009.
- REISINGER, Don. "Can Google Android Beat Windows 7 in the Netbook Market?". *eWeek*, 5 jun. 2009.
- ROGOW, Bruce. "Tracking Core Assets". *Optimize Magazine*, abr. 2006.
- SANDLER, Kathy. "The Future of Touch". *The Wall Street Journal*, 2 jun. 2009.
- SUNGARD AVAILABILITY SERVICES. "Sungard Availability Services' Responsible Computing Project Powers Worldwide Volunteer Research". 31 ago. 2009.
- TOIGO, Jon. "Storage in the Cloud". *Information Week*, 11 maio 2009.
- VANCE, Ashlee; RICHTEL, Matt. "Light and Cheap, Netbooks Are Poised to Reshape PC Industry". *The New York Times*, 2 abr. 2009.
- VANCE, Ashlee. "Microsoft Office 2010 Starts Ascension to the Cloud". *The New York Times*, 13 jul. 2009.
- VANCE, Ashlee. "PC Touch Screens Move Ahead". *The New York Times*, 2 jun. 2009.
- VASCELLARO, Jessica. "Google Strives to Help Online Software Catch Up". *The Wall Street Journal*, 15 jul. 2009.
- WALSH, Lawrence. "Outsourcing: A Means of Business Enablement". *Baseline*, maio 2008.
- WEIER, Mary Hayes. "Business Gone Mobile". *Information Week*, 30 mar. 2009.
- WEILL, Peter; SUBRAMANI, Mani; BROADBENT, Marianne. "Building IT Infrastructure for Strategic Agility". *Sloan Management Review*, v. 44, n. 1, primavera 2002.
- WILLIAMS, Mark. "The Digital Utility". *Technology Review*, mar./abr. 2008.

Fundamentos da inteligência de negócios: gestão da informação e de bancos de dados

Capítulo

5

OBJETIVOS DE ESTUDO

Ao concluir este capítulo, você será capaz de responder às seguintes perguntas:

1. Como um banco de dados relacional organiza os dados e compara essa abordagem ao banco de dados orientado a objetos?
2. Quais os princípios de um sistema de gestão de banco de dados?
3. Quais as principais ferramentas e tecnologias para extrair informações de bancos de dados e melhorar o desempenho da empresa e a tomada de decisão?
4. Qual o papel da política de informação e da gestão de dados na organização dos recursos de dados empresariais?
5. Por que é importante para a empresa garantir a qualidade dos dados?

PLANO DO CAPÍTULO

Caso de abertura: R. R. Donnelley tenta controlar seus dados

A abordagem de banco de dados para gestão de dados

Sistemas de gestão de banco de dados

Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa

Gestão dos recursos de dados

Projetos práticos em SIG

Resolvendo problemas organizacionais — Problemas com o banco de dados sobre terroristas

A R.R. DONNELLEY TENTA CONTROLAR SEUS DADOS

É possível que você esteja utilizando um produto R. R. Donnelley neste momento. A empresa, com sede em Chicago, é um gigante na área de serviços e impressões comerciais, oferecendo serviços de impressão, formulários e etiquetas, mala direta e outros. A recente expansão dos negócios foi recheada por uma série de aquisições, incluindo a tipografia comercial Moore Wallace, em 2005, e a empresa de gestão de impressão e cadeia de suprimentos Banta, em 2007. A receita da R. R. Donnelley subiu de 2,4 bilhões de dólares, em 2003, para 9,8 bilhões de dólares em 2009.

Entretanto, todo o crescimento criou desafios para a gerência de informações. Cada empresa adquirida contava com seus próprios sistemas e seu próprio grupo de dados de clientes, fornecedores e produtos. Vindos de tantas fontes diferentes, os dados costumavam ser inconsistentes, duplicados ou incompletos. Por exemplo, cada unidade diferente da empresa pode assumir um significado para a entidade 'cliente'. Uma pode definir como um local específico de faturamento, outra pode defini-lo como a entidade-mãe legal de uma empresa. Donnelley precisava usar processos manuais demorados para conciliar os dados armazenados em diferentes sistemas a fim de obter uma imagem clara de cada cliente, visto que eles podem estar negociando com diversas unidades da empresa. Essas condições acentuavam as ineficiências e os custos.

R. R. Donnelley tornou-se tão grande que era impraticável armazenar as informações de todas as suas unidades em um único sistema. Precisava, ainda, de um conjunto de dados único para toda a empresa que fosse preciso e consistente. Para resolver esse problema, R. R. Donnelley recorreu à gestão de dados mestres (MDM – *master data management*), que busca garantir que uma empresa não utilize múltiplas versões dos mesmos dados em diferentes partes de suas opera-

ções, através da consolidação de registros díspares em um único arquivo mestre autenticado. Uma vez que o arquivo mestre esteja disponível, empregados e aplicações acessam uma visão única consolidada dos dados. Esse recurso é especialmente útil para organizações como a Donnelley, com problemas na integração dos dados devido a fusões e aquisições.

A implantação de um MDM é um processo de múltiplas etapas que inclui análise de processos de negócios; limpeza, consolidação e reconciliação de dados, e migração destes para um arquivo mestre de todos os dados da empresa. As organizações devem identificar o grupo 'proprietário' de cada um dos dados e responsável por resolver definições inconsistentes de dados e outras discrepâncias.

A Donnelley lançou seu programa de MDM no final de 2005 e começou a criar um conjunto único de identificadores para os dados de seus clientes e fornecedores. Ela optou por um modelo de registros utilizando o Data Hub da Purisma — no qual os dados dos clientes continuam a residir no sistema onde foram criados, mas são registrados em um centralizador (*hub*) e são relacionados de modo que as aplicações possam encontrá-los. Os dados em seus sistemas de origem permanecem intocados.

Quase um ano mais tarde, a Donnelley criou seu Armazenamento de Dados Mestre de Clientes (*customer master data store*), que integra os dados de inúmeros sistemas das aquisições da empresa. Os dados ultrapassados, incompletos ou formatados incorretamente são corrigidos ou eliminados. Um registro aponta para o local onde os dados-fonte estão armazenados. Ao dispor de um conjunto de dados consistentes de toda a empresa, com definições e padrões comuns, a gerência pode descobrir facilmente que tipo de negócio possui com determinado cliente, bem como seu volume, para



identificar a principal clientela e as oportunidades de vendas. Atualmente, quando a Donnelley compra uma empresa, pode facilmente descobrir quais os clientes duplicados.

Fontes: John McCormick, "Mastering data at R.R. Donnelley". *Information Management Magazine*, mar. 2009; Loraine Lawson, "Are you really ready to buy a master data management tool?". *IT BusinessEdge*, 25 jun. 2009; e www.purisma.com, acesso em 26 jun. 2009.

A experiência da R. R. Donnelley ilustra a importância da gestão de dados para os negócios. A empresa viveu um crescimento fenomenal, primeiramente pelas aquisições. Seu desempenho, entretanto, depende do que consegue ou não fazer com seus dados. A forma como as empresas armazenam, organizam e gerenciam suas informações causa um impacto tremendo sobre a eficiência empresarial.

A figura de abertura do caso nos chama a atenção para pontos importantes levantados pelo caso e pelo capítulo. A gerência decidiu que a empresa precisava centralizar a gestão dos dados da empresa. Informações sobre clientes, fornecedores, produtos e outras entidades importantes foram armazenadas em diferentes sistemas e arquivos nos quais não podiam ser facilmente recuperadas e analisadas. Eram sempre redundantes e inconsistentes, limitando sua utilidade. A gerência não conseguia obter uma visão geral de todos os clientes de todas as suas aquisições para comercializar seus produtos e serviços e oferecer melhores suporte e atendimento.

No passado, a R. R. Donnelley utilizou processos manuais para agrupar seus dados inconsistentes e redundantes e gerenciar as informações de uma perspectiva que abrangesse toda a empresa. Essa solução já não era mais viável à medida que começou a crescer. Uma solução mais apropriada era identificar, consolidar, limpar e padronizar todos os dados em um único registro de gestão de dados mestres. Além de utilizar a tecnologia apropriada, a Donnelley precisou corrigir e reorganizar os dados em um formato padrão, além de estabelecer regras, responsabilidades e procedimentos para atualização e uso dos dados.

Um sistema de gestão de dados mestres ajuda a R. R. Donnelley a aumentar a lucratividade facilitando a identificação de clientes e oportunidades de venda. Ele também aumenta a eficiência operacional e a tomada de decisões ao disponibilizar informações mais precisas e completas e ao reduzir o tempo necessário para ajustar os dados redundantes e inconsistentes.

A abordagem de banco de dados para gestão de dados

Um **banco de dados** é um conjunto de arquivos relacionados entre si com registros sobre pessoas, lugares ou coisas. Um dos bancos de dados mais bem-sucedidos da história moderna é a lista telefônica, conjunto de registros referentes a pessoas físicas e jurídicas que possuem telefone. Ela fornece quatro tipos de informação sobre cada usuário de telefone: sobrenome, primeiro nome, endereço e número de telefone. Traz também informações sobre empresas e categorias de empresa, tais como lojas de automóveis ou fornecedores de material hidráulico. A lista telefônica extrai essas informações de um banco de dados com arquivos de clientes, classificações de empresa, códigos de área e regiões geográficas.

Antes do desenvolvimento dos bancos de dados digitais, as empresas usavam grandes arquivos de metal, cheios de arquivos de papel, para armazenar informações referentes a transações, clientes, fornecedores, estoque e funcionários. Costumavam usar também listas laboriosamente coletadas e datilografadas para resumir as informações dos arquivos de papel. Ainda podemos encontrar bancos de dados manuais, feitos em papel, em muitos consultórios médicos, onde os registros dos pacientes são armazenados em centenas de fichas.

Nem é preciso dizer que os bancos de dados em papel são extremamente ineficientes e caros de manter, muitas vezes contêm dados incorretos, são lentos e dificultam o acesso imediato aos dados. Além disso, são extremamente inflexíveis. Por exemplo, seria difícil para um consultório médico com dados em papel combinar seus arquivos sobre prescrições com seus arquivos sobre pacientes, para produzir uma lista de todas as pessoas a quem foi prescrito determinado remédio. Para um banco de dados moderno, digital, isso seria fácil. De fato, uma característica imbatível dos bancos de dados é a capacidade de relacionar rapidamente um conjunto de arquivos a outro.

Entidades e atributos

Ao pensar nos dados para sua empresa e em como gerenciá-los, por onde começaria? Se está abrindo sua empresa ou já gerencia uma, precisa identificar os dados necessários à administração do negócio. Normalmente, usará dados referentes a determinados tipos de informação, tais como clientes, fornecedores, funcionários, pedidos, produtos, expedidores e talvez peças. Cada uma dessas categorias genéricas que representa uma pessoa, um lugar ou uma coisa sobre a qual você armazena e mantém informações é denominada **entidade**. Cada entidade tem características específicas, chamadas **atributos**. Por exemplo, a entidade FORNECEDOR tem atributos específicos, como nome e endereço de fornecedor, o que engloba rua, cidade, estado e código postal. A entidade PEÇA normalmente tem atributos como descrição de peça, preço de cada peça (preço unitário) e fornecedor que produziu a peça.

Organizando dados em um banco de dados relacional

Se você armazenasse essas informações em arquivos de papel, provavelmente teria um arquivo para cada entidade e seus atributos. Em um sistema de informação, o banco de dados organiza as informações de maneira muito parecida, agrupando dados relacionados entre si. O **banco de dados relacional** é o tipo de banco de dados mais comum atualmente. Bancos de dados relacionais organizam os dados em tabelas bidimensionais (denominadas relações) com colunas e linhas. Cada tabela contém dados referentes a uma entidade e seus atributos. Na maioria dos casos, monta-se uma tabela para cada entidade do negócio. Assim, no nível mais básico, temos uma tabela para clientes, uma para fornecedores, outra para peças em estoque, outra para funcionários e outra para transações de venda.

Vamos examinar como um banco de dados relacional organizaria dados a respeito de fornecedores e peças. Tome como exemplo a tabela FORNECEDOR, ilustrada na Figura 5.1. Ela consiste em uma grade de colunas e linhas de dados. Cada dado específico a respeito de um fornecedor, como nome, endereço, cidade, estado e código postal do fornecedor, é armazenado como um **campo** individual dentro da tabela FORNECEDOR. Cada campo representa um atributo da entidade FORNECEDOR. Nos bancos de dados relacionais, os campos também são chamados de **colunas**.

A informação específica sobre um único fornecedor armazenada em uma tabela é denominada linha. As linhas também são muitas vezes chamadas de **registros** ou, em termos bem técnicos, **tuplas**.

Observe que nessa tabela há um campo para Número_Fornecedor. Esse campo, denominado **campo-chave**, atribui uma identificação exclusiva a cada registro, permitindo que seja recuperado, atualizado e ordenado. Cada tabela de um banco de dados relacional tem um campo designado **chave primária**. Esse campo-chave é o identificador exclusivo para todas as informações em qualquer linha da tabela, e essa chave primária não pode ser duplicada.

Figura 5.1

Tabela de banco de dados relacional

Um banco de dados relacional organiza os dados na forma de tabelas bidimensionais. Temos aqui uma tabela para a entidade FORNECEDOR e vemos como ela representa a entidade e seus atributos. Número_Fornecedor é o campo-chave.

FORNECEDOR

Colunas (Atributos, Campos)

Número_Fornecedor	Nome_Fornecedor	Rua_Fornecedor	Cidade_Fornecedor	Estado_Fornecedor	CEP_Fornecedor
8259	CBM Inc.	74 5 th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Composites	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Linhas (Registros, Tuplas)

Campo-chave (chave primária)

Poderíamos usar o nome do fornecedor como campo-chave. No entanto, se dois fornecedores diferentes tiverem o mesmo nome (o que de fato ocorre de vez em quando), o nome do fornecedor não será uma identificação exclusiva para cada um deles, portanto é necessário designar um campo identificador especial a esse propósito. Por exemplo, se houver dois fornecedores, ambos chamados “CBM”, um localizado em Dayton e outro em St. Louis, será fácil confundi-los. No entanto, se cada um tiver um Número_Fornecedor exclusivo, a confusão será evitada.

Vemos também que a informação postal foi separada em quatro campos: Rua_Fornecedor, Cidade_Fornecedor, Estado_Fornecedor e CEP_Fornecedor. Os dados são separados nas menores unidades que desejamos acessar separadamente, para que seja fácil selecionar na tabela somente as linhas coincidentes com o conteúdo de um único campo — todos os fornecedores em Ohio (OH), por exemplo. As linhas de dados também podem ser ordenadas pelo conteúdo do campo Estado_Fornecedor; dessa maneira, teremos uma lista de fornecedores por estado, independentemente da cidade.

Até agora, a tabela FORNECEDOR não possui nenhuma informação sobre as peças que um fornecedor particular provê à sua empresa. PEÇA é uma entidade separada de FORNECEDOR, logo os campos com informações sobre peças devem ser armazenados em uma tabela PEÇA independente (veja a Figura 5.2).

Por que não guardamos as informações sobre peças na mesma tabela dos fornecedores? Se fizermos isso, cada linha da tabela conterá atributos referentes a PEÇA e atributos referentes a FORNECEDOR. Como um fornecedor pode abastecer mais de uma peça, serão necessárias muitas linhas em cada fornecedor para mostrar todas as peças que ele provê. Teremos, assim, grande quantidade de dados redundantes sobre fornecedores, e será difícil buscar as informações sobre uma peça particular, porque não saberemos se essa peça é a primeira ou a décima quinta no registro do fornecedor. Portanto, uma tabela separada, PEÇA, deve ser criada para armazenar esses três campos e resolver o problema.

A tabela PEÇA também deve conter outro campo, o Número_Fornecedor, para que você saiba o fornecedor de cada peça. Não será necessário repetir todas as informações sobre um fornecedor em cada registro de peça porque, tendo um campo Número_Fornecedor na tabela PEÇA, você pode ‘pesquisar’ os dados nos campos da tabela FORNECEDOR.

Observe que Número_Fornecedor aparece nas tabelas FORNECEDOR e PEÇA. Na tabela FORNECEDOR, Número_Fornecedor é a chave primária. Quando esse campo aparece na tabela PEÇA, porém, é denominado **chave estrangeira** e, essencialmente, é um campo de pesquisa usado para obter dados a respeito do fornecedor de uma peça específica. Observe que a tabela PEÇA tem sua própria chave primária, Número_Peça, que atribui uma identificação exclusiva a cada peça. Essa chave não é usada para vincular PEÇA a FORNECEDOR, mas poderia ser usada para vincular PEÇA a outra entidade qualquer.

Figura 5.2

A tabela PEÇA

Dados da entidade PEÇA têm sua própria tabela individual. Número_Peça é a chave primária e Número_Fornecedor é a chave estrangeira, que permite aos usuários achar, na tabela FORNECEDOR, informações correlatas sobre o fornecedor de cada peça.

PEÇA

Número_Peça	Descrição_Peça	Preço_Unitário	Número_Fornecedor
137	Trinco da porta	22,00	8259
145	Retrovisor externo	12,00	8444
150	Vedação da porta	6,00	8263
152	Trava da porta	31,00	8259
155	Compressor	54,00	8261
178	Maçaneta da porta	10,00	8259

Chave primária

Chave estrangeira

Ao organizar dados em tabelas, é importante cuidar para que todos os atributos de uma entidade específica se apliquem apenas a essa entidade. Se colocar o endereço do fornecedor no registro PEÇA, essa informação na verdade não estará relacionada apenas à PEÇA; estará relacionada tanto a PEÇA quanto a FORNECEDOR. Assim, se o endereço do fornecedor mudar, será necessário alterar os dados em cada registro PEÇA, em vez de fazê-lo uma única vez no registro FORNECEDOR.

Estabelecendo relacionamentos

Agora que separamos nossos dados na tabela FORNECEDOR e na tabela PEÇA, precisamos ter certeza de que entendemos a relação entre as duas tabelas. Uma representação esquemática denominada **diagrama entidade/relacionamento** é usada para esclarecer relações entre as tabelas de um banco de dados relacional. A maior utilidade desse diagrama é dizer de que maneira duas tabelas se relacionam. Em um banco de dados relacional, as tabelas podem ter relacionamentos um-para-um, um-para-muitos ou muitos-para-muitos.

Um exemplo de relação um-para-um pode ser uma situação em que um sistema de recursos humanos precisa armazenar dados confidenciais sobre os funcionários. Podem-se armazenar dados tais como nome do funcionário, data de nascimento, endereço e cargo em uma tabela, e os dados confidenciais sobre o funcionário, como salário ou benefícios, em outra. Essas duas tabelas pertencentes a um único funcionário teriam uma relação um-para-um, porque cada registro na tabela FUNCIONÁRIO, com os dados básicos do funcionário, teria apenas um registro correlato na tabela que armazena os dados confidenciais.

A relação entre as entidades FORNECEDOR e PEÇA em nosso banco de dados é uma relação do tipo um-para-muitos: cada fornecedor pode prover mais de uma peça, mas cada peça tem apenas um único fornecedor. Para cada registro na tabela FORNECEDOR, pode haver muitos registros correlatos na tabela PEÇA.

A Figura 5.3 ilustra como o diagrama entidade/relacionamento descreveria essa relação um-para-muitos. As caixas representam as entidades, enquanto as linhas que ligam uma caixa a outra representam relacionamentos. Uma linha que termina com dois tracinhos representa um relacionamento um-para-um. Uma linha que termina com um tridente e um tracinho atrás indica uma relação um-para-muitos. A Figura 5.3 mostra que cada PEÇA tem apenas um FORNECEDOR, mas muitas PEÇAS podem ser fornecidas pelo mesmo FORNECEDOR.

Teremos também uma relação um-para-muitos se quisermos acrescentar uma tabela referente a pedidos no nosso banco de dados, porque um fornecedor atende a muitos pedidos. A tabela PEDIDO conteria apenas o Número_Pedido e a Data_Pedido. A Figura 5.4 ilustra um relatório mostrando um pedido de várias peças de determinado fornecedor. Como se vê, as informações no canto superior direito vêm da tabela PEDIDO. Os itens efetivamente pedidos são listados na parte inferior do relatório.

Como um pedido pode englobar muitas peças de um mesmo fornecedor, e uma única peça pode ser pedida muitas vezes em diferentes pedidos, isso cria uma relação muitos-para-muitos entre as tabelas PEÇA e PEDIDO. Sempre que existe uma relação muitos-para-muitos entre duas tabelas, é necessário combinar essas duas tabelas em uma única que reúna tais informações. Para tanto, podemos criar uma tabela separada para um item no pedido. Essa tabela muitas vezes é denominada *tabela concatenada* ou *relação de intersecção*. A tabela concatenada contém apenas três campos (sendo dois deles, Número_Pedido e Número_Peça, utilizados para concatenar as tabelas PEDIDO e PEÇA) e a Quantidade_Peça. Como se vê, as informações no canto inferior esquerdo do relatório vêm da tabela ITEM.

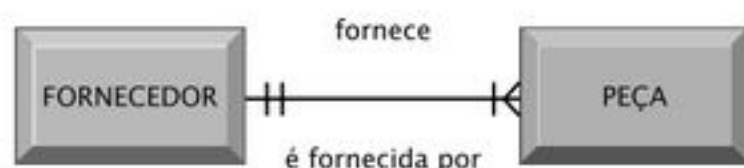


Figura 5.3

Um diagrama entidade/relacionamento básico

Este diagrama mostra a relação entre as entidades FORNECEDOR e PEÇA.

Figura 5.4

Amostra de relatório de pedido

As áreas sombreadas mostram quais dados vêm das tabelas FORNECEDOR, PEDIDO e ITEM. O banco de dados não mantém dados sobre Preço Total nem Pedido Total porque eles podem ser derivados a partir de outros dados das tabelas.

Número_Pedido	Número_Peça	Quantidade_Peça	Nome_Peça	Preço_Unitário	Preço Total
3502	137	10	Trinco da porta	22,00	220,00
3502	152	20	Trava da porta	31,00	620,00
3502	178	5	Maçaneta da porta	10,00	50,00
Pedido Total:					890,00

Número do pedido: 3502
Data do pedido: 15/01/2006

Número do fornecedor: 8259
Nome do fornecedor: CBM Inc.
Endereço do fornecedor: 74 5th Avenue, Dayton, OH 45220

Ficaremos, então, com um total de quatro tabelas no nosso banco de dados. A Figura 5.5 ilustra o conjunto final de tabelas, enquanto a Figura 5.6 mostra o diagrama entidade/relacionamento para esse conjunto de tabelas. Observe que a tabela PEDIDO não contém dados sobre o preço total de cada item porque esse valor pode ser calculado multiplicando-se Preço_Unitário por Quantidade_Peça. Esse dado pode, enfim, ser *derivado* quando necessário a partir de informações já existentes nas tabelas PEÇA e ITEM. O Pedido Total, outro campo derivado, é calculado pela soma do preço total de cada item encomendado.

O processo de simplificar grupos complexos de dados a fim de minimizar elementos de dados redundantes e relações muitos-para-muitos inadequadas, e ao mesmo tempo aumentar a estabilidade e a flexibilidade, chama-se **normalização**. Um banco de dados adequadamente projetado e normalizado será fácil de manter e terá pouca duplicação de dados.

Figura 5.5

Projeto final do banco de dados com amostras de registros

O projeto final do banco de dados referente a fornecedores, peças e pedidos tem quatro tabelas. A tabela ITEM é uma tabela concatenada que elimina a relação muitos-para-muitos entre PEDIDO e PEÇA.

PEÇA			
Número_Peça	Descrição_Peça	Preço_Unitário	Número_Fornecedor
137	Trinco da porta	22,00	8259
145	Retrovisor externo	12,00	8444
150	Vedação da porta	6,00	8263
152	Trava da porta	31,00	8259
155	Compressor	54,00	8261
178	Maçaneta da porta	10,00	8259

ITEM			PEDIDO	
Número_Pedido	Número_Peça	Quantidade_Peça	Número_Pedido	Data_Pedido
3502	137	10	3502	1/15/2006
3502	152	20	3502	1/15/2006
3502	178	5	3502	1/15/2006

FORNECEDOR					
Número_Fornecedor	Nome_Fornecedor	Rua_Fornecedor	Cidade_Fornecedor	Estado_Fornecedor	CEP_Fornecedor
8259	CBM Inc.	74 5 th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Components	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344



Figura 5.6

Diagrama entidade/relacionamento para o banco de dados de quatro tabelas

Este diagrama mostra o relacionamento entre as entidades FORNECEDOR, PEÇA, ITEM e PEDIDO.

Os sistemas de bancos de dados relacionais tentam impor regras de **integridade referencial** para garantir que os relacionamentos entre as tabelas permaneçam consistentes. Quando uma tabela possui uma chave estrangeira que aponta para outra tabela, você pode só conseguir incluir um registro na tabela com a chave estrangeira se houver um correspondente na tabela relacionada. No banco de dados que acabamos de criar, a chave estrangeira `Número_Fornecedor` relaciona a tabela `PEÇA` à tabela `FORNECEDOR`. Podemos não conseguir incluir um registro na tabela `PEÇA` para uma peça fornecida pelo fornecedor 8266 até que exista um registro correspondente na tabela `FORNECEDOR` para o fornecedor número 8266. Se excluirmos o fornecedor número 8266 da tabela `FORNECEDOR`, precisamos também excluir o registro correspondente na tabela `PEÇA`. Em outras palavras, não devem existir peças para fornecedores fantasmas!

O exemplo fornecido aqui é muito simples. Mesmo em uma empresa bem pequena você terá tabelas para outras entidades importantes, tais como clientes, expedidores e funcionários. Uma grande corporação talvez precise manter um banco de dados com milhares de entidades (tabelas). O que importa para qualquer empresa, pequena ou grande, é ter um bom modelo de dados, que abranja todas as suas entidades e o relacionamento entre elas, e esteja organizado para minimizar a redundância, maximizar a precisão e deixar os dados facilmente acessíveis a relatórios e análises.

Nunca é demais enfatizar: se a empresa não trabalhar com o modelo de dados correto, o sistema não será capaz de atendê-la corretamente. Seus sistemas não funcionarão tão bem quanto poderiam, porque trabalharão com dados que podem ser imprecisos, incompletos ou difíceis de recuperar. Entender os dados de uma organização e saber como eles devem ser representados em um banco de dados é talvez a lição mais importante que se pode aprender.

Por exemplo, a Famous Footwear, cadeia de lojas de calçados com mais de 800 filiais em 49 estados norte-americanos, não conseguia cumprir sua meta de ter à venda “o estilo certo de calçado, na loja certa, pelo preço certo” porque seu banco de dados não estava projetado para ajustar o estoque das lojas com a velocidade necessária. A empresa tinha um banco de dados relacional da Oracle, que rodava em um computador de médio porte IBM AS/400, mas o banco de dados estava projetado fundamentalmente para produzir relatórios administrativos padrão, e não para reagir a mudanças no mercado. A gerência não conseguia obter dados precisos sobre itens específicos estocados em cada loja. Para resolver o problema, desenvolveu um novo banco de dados, no qual as informações sobre vendas e estoques foram mais bem organizadas para análise e gestão de estoque.

Sistemas de gestão de banco de dados

Agora que começou a criar os arquivos e identificar os dados necessários para administrar sua empresa, você precisará de um sistema de gestão de banco de dados para ajudá-lo a gerenciar e usar esses dados. Um **sistema de gestão de banco de dados (DBMS – database management system)** é um software específico usado para criar, armazenar, organizar e acessar dados a partir de um banco de dados. O Microsoft Access é um DBMS para computadores pessoais, enquanto o DB2, o Oracle e o Microsoft SQL Server são DBMSs para grandes mainframes e computadores de médio porte. O MySQL é um popular DBMS de código aberto, enquanto o Oracle Database Lite é um DBMS para pequenos dispositivos digitais de mão. Todos esses produtos são DBMSs relacionais que suportam um banco de dados relacional.

O DBMS livra o programador ou o usuário final da tarefa de entender onde e como os dados estão realmente armazenados, separando as visões lógica e física dos dados. A *visão lógica* apresenta os dados tais como seriam vistos por usuários ou especialistas da empresa,

ao passo que a *visão física* mostra como eles estão realmente organizados e estruturados nos meios de armazenamento físico, como um disco rígido.

O software de gestão de banco de dados disponibiliza o banco físico para as diferentes visões lógicas solicitadas pelos usuários. Por exemplo, no caso do banco de dados de recursos humanos ilustrado na Figura 5.7, um especialista em benefícios pode requerer uma visão que consista no nome do funcionário, em seu número da Previdência Social e na cobertura de seu plano de saúde. Já um funcionário do departamento responsável pela folha de pagamentos pode precisar de dados como o nome do funcionário, seu número da Previdência Social, seu salário bruto e seu salário líquido. Os dados para todas essas visões são armazenados em um único banco de dados, no qual eles podem ser administrados mais facilmente pela organização.

Operações de um DBMS relacional

Em um banco de dados relacional, as tabelas podem ser facilmente combinadas para apresentar dados solicitados por usuários, desde que compartilhem um elemento de dado em comum. Vamos retornar ao banco de dados que montamos anteriormente com as tabelas PEÇA e FORNECEDOR, ilustradas nas figuras 5.1 e 5.2.

Como demonstração, suponha que se queira descobrir nesse banco de dados os nomes dos fornecedores que podem prover a peça número 137 ou 150. Você precisaria de informações de duas tabelas: FORNECEDOR e PEÇA. Observe que essas duas tabelas compartilham um elemento de dado: Número_Fornecedor.

No banco de dados relacional, três operações básicas, conforme mostrado na Figura 5.8, são utilizadas para desenvolver conjuntos de dados úteis: *select*, *project* e *join* (selecionar, projetar e vincular). A operação *select* cria um subconjunto formado de todos os registros (linhas) da tabela que obedecem a critérios estabelecidos. Em outras palavras, a instrução *select* cria um subconjunto de linhas que atendem a determinado critério. Em nosso exemplo, queremos selecionar registros (linhas) da tabela PEÇA em que o Número_Peça é igual a 137 ou 150. A operação *join* combina as tabelas relacionais para fornecer ao usuário mais informações do que aquelas à disposição nas tabelas individuais. Em nosso exemplo, queremos vincular a tabela PEÇA já reduzida (apenas as peças de números 137 ou 150 serão apresentadas) com a tabela FORNECEDOR, para originar uma única e nova tabela.

A operação *project* cria um subconjunto composto de colunas de uma tabela que permite ao usuário criar novas tabelas contendo apenas as informações requisitadas. Em nosso exemplo, queremos extrair da nova tabela somente as seguintes colunas: Número_Peça, Descrição_Peça, Número_Fornecedor e Nome_Fornecedor (veja a Figura 5.8).

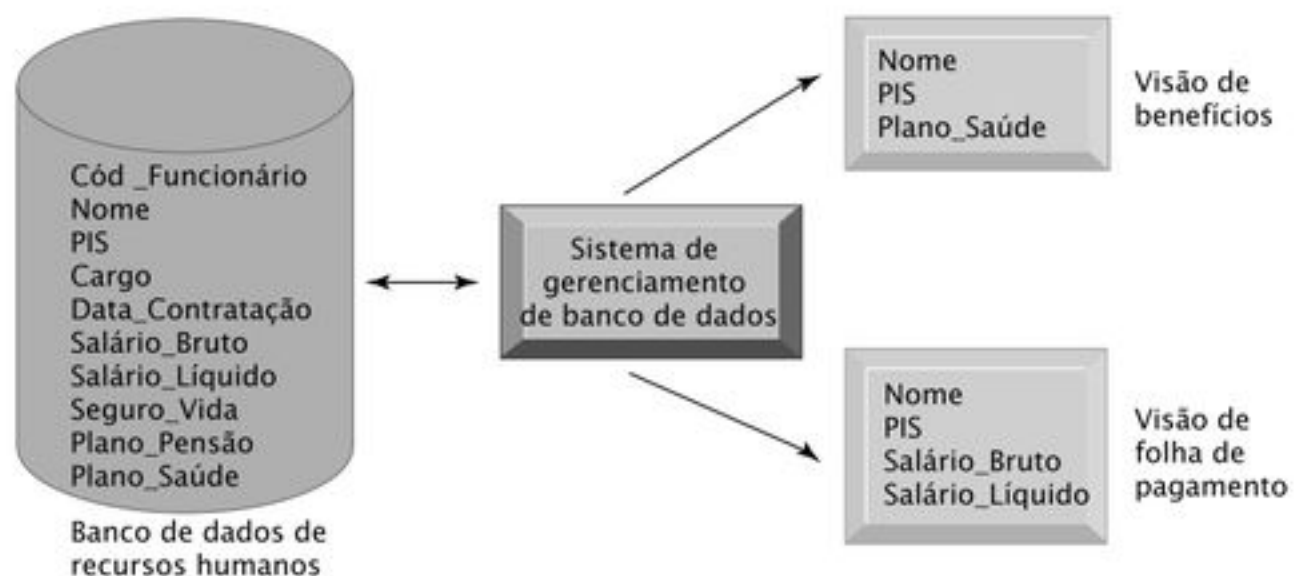
Figura 5.7

Banco de dados de recursos humanos com múltiplas visões

Um único banco de dados de recursos humanos oferece muitas visões dos dados, conforme as necessidades de informação do usuário. Exemplificamos aqui duas visões possíveis, uma que interessa a um especialista em benefícios, e outra que interessa a um funcionário do departamento responsável pela folha de pagamentos da empresa.

Recursos dos sistemas de gestão de banco de dados

Um DBMS possui recursos e ferramentas para organizar, administrar e acessar os dados do banco de dados. Os mais importantes são o recurso de definição de dados, o dicionário de dados e a linguagem de manipulação de dados.



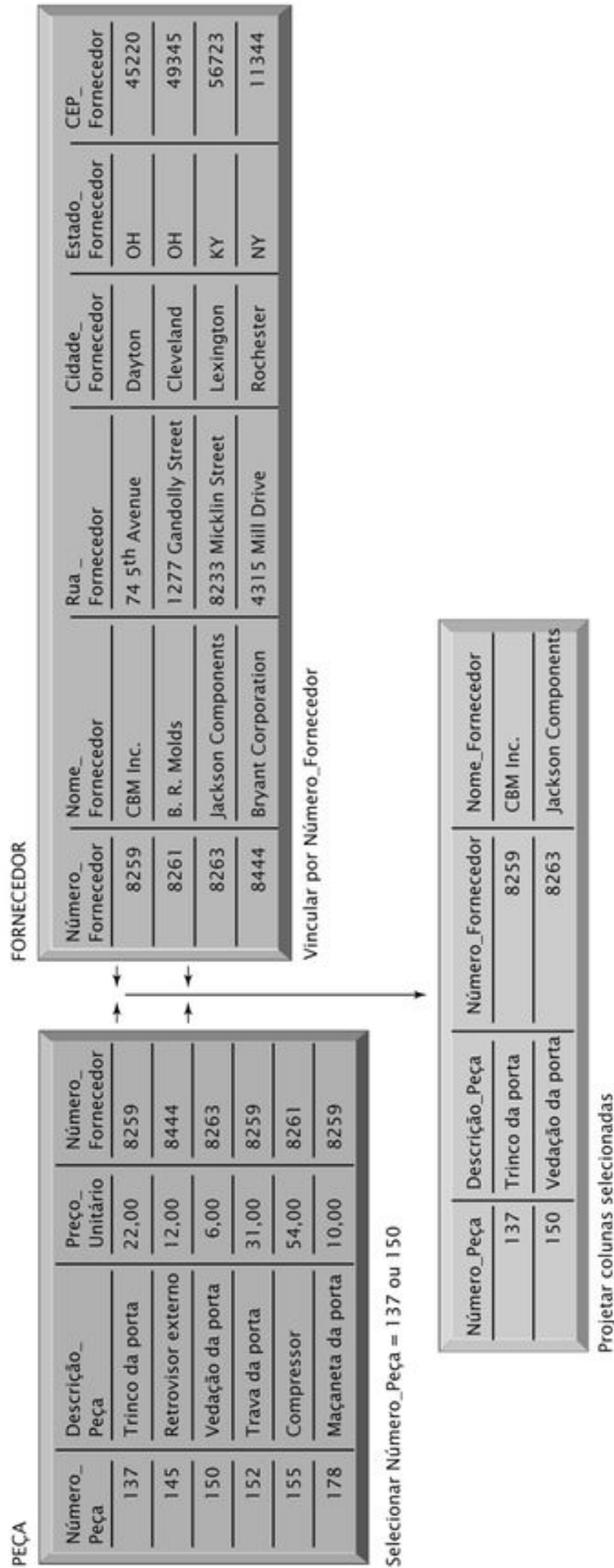


Figura 5.8

As três operações básicas de um DBMS relacional

As operações select, project e join (selecionar, projetar e vincular) permitem que dados de duas tabelas diferentes sejam combinados e que apenas os atributos selecionados sejam apresentados.

Os DBMSs têm um recurso de **definição de dados** para especificar a estrutura do conteúdo do banco de dados. Esse recurso pode ser usado para criar tabelas de banco de dados e para definir as características dos campos em cada tabela. Tais informações sobre o banco de dados costumam ser documentadas no **dicionário de dados**, um arquivo, manual ou automatizado, que armazena as definições dos elementos de dado e suas características. O Microsoft Access oferece um recurso rudimentar de dicionário de dados pré-instalado, que apresenta informações quanto a nome, descrição, tamanho, tipo e formato de cada campo em uma tabela (veja a Figura 5.9). Em bancos de dados de grandes corporações, os dicionários de dados podem contemplar outras informações, como uso; proprietário (quem, dentro da organização, é responsável pela manutenção dos dados); autorização; segurança; e os indivíduos, as funções empresariais, os programas e os relatórios que usam cada elemento de dado.

Consultas e relatórios

Os DBMSs trazem ferramentas para acesso e manipulação de informações em bancos de dados. A maioria dos DBMSs possui uma linguagem especializada, a **linguagem de manipulação de dados**, usada para acrescentar, alterar, apagar e recuperar os dados do banco de dados. Essa linguagem contém comandos que permitem aos usuários finais e especialistas em programação extrair dados do banco de dados para satisfazer requisições de informações e desenvolver aplicações. A linguagem de manipulação de dados mais usada atualmente é a **linguagem estruturada de consulta** ou **SQL (Structured Query Language)**. A Figura 5.10 ilustra a consulta SQL que produziria a nova tabela de resultados da Figura 5.8.

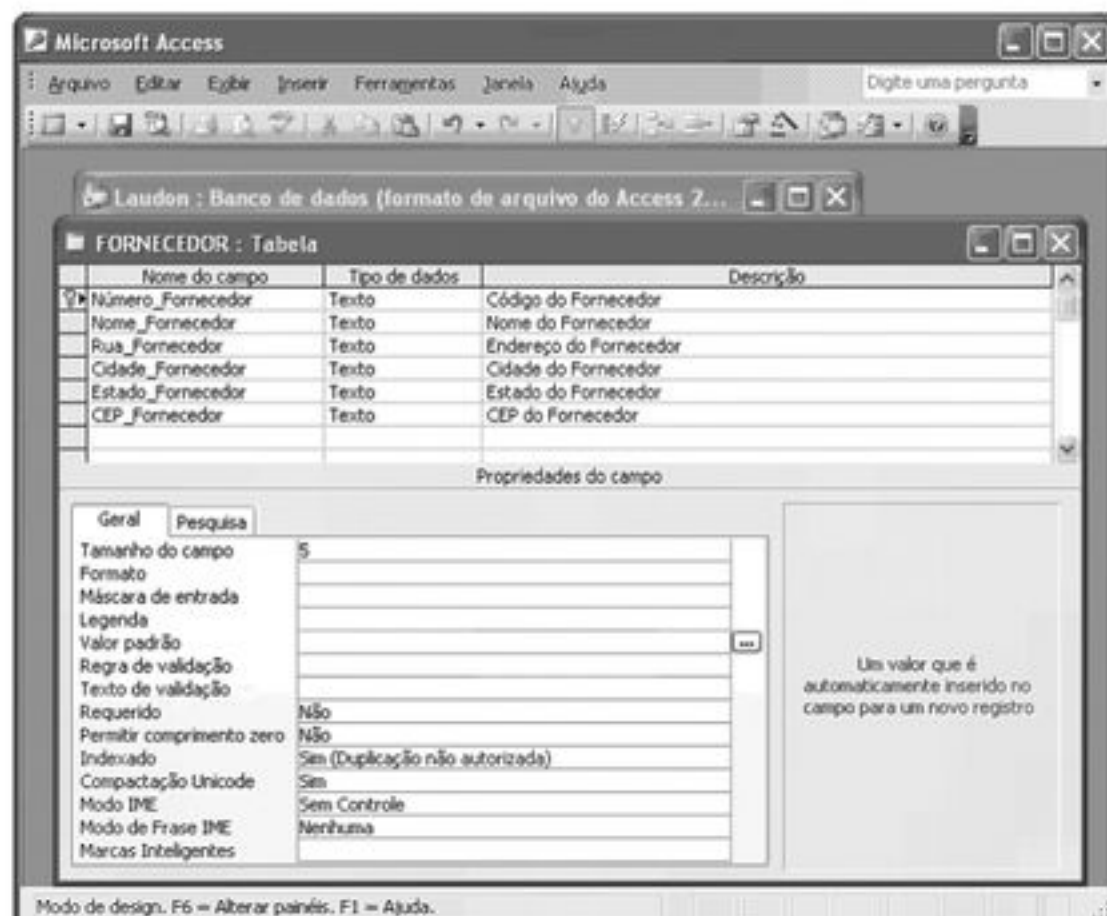
Usuários de DBMS para computadores de grande e médio portes, tais como DB2, Oracle ou SQL Server, costumam usar SQL para recuperar as informações do banco de dados. O Microsoft Access também usa SQL, mas fornece seu próprio conjunto de ferramentas amigáveis para consultar bancos de dados e organizar dados em relatórios mais refinados.

O Microsoft Access dispõe de recursos que permitem ao usuário criar consultas. Para tanto, o usuário deve identificar as tabelas e os campos que deseja, bem como os resultados e, em seguida, selecionar as linhas do banco de dados que atendem a determinados critérios. Essas ações são então traduzidas em comandos SQL. A Figura 5.11 mostra como a mesma consulta em SQL para selecionar peças e fornecedores, ilustrada na Figura 5.10, seria construída usando Microsoft Access.

Figura 5.9

Mudou a versão do Access no qual o banco de dados foi criado, mas as informações continuam as mesmas. Aparência do dicionário de dados do Access

O Microsoft Access traz um recurso rudimentar de dicionário de dados que apresenta informações quanto a tamanho, formato e outras características de cada campo em um banco de dados. Aqui, vemos as informações mantidas na tabela FORNECEDOR. O pequeno ícone representando uma chave, à esquerda de Número_Fornecedor, indica que se trata de um campo-chave.



```
SELECT PEÇA.Número_Peça, PEÇA.Descrição_Peça,
FORNECEDOR.Número_Fornecedor, FORNECEDOR.Nome_Fornecedor
FROM PEÇA, FORNECEDOR
WHERE PEÇA.Número_Fornecedor = FORNECEDOR.Número_Fornecedor AND
Número_Peça = 137 OR Número_Peça = 150;
```

Normalmente, os DBMSs incluem recursos para geração de relatórios, de maneira que os dados de interesse possam ser mostrados em um formato mais refinado e estruturado do que uma consulta simples. O Crystal Reports é um popular gerador de relatórios para grandes DBMSs corporativos, embora também possa ser usado com o Access.

O Access conta ainda com recursos para desenvolvimento de aplicativos para PC. Entre eles, ferramentas para criar telas de entrada de dados, relatórios e para desenvolver a lógica a ser aplicada ao processamento de transações. Esses recursos são normalmente usados por especialistas em sistemas de informação.

Bancos de dados orientados a objeto

Muitas aplicações exigem, hoje, bancos de dados que possam armazenar e recuperar não somente números e caracteres estruturados, como também desenhos, imagens, fotografias, voz e vídeo com movimento. O DBMS convencional, projetado para organizar dados estruturados em linhas e colunas, não é adequado para tratar aplicações baseadas em recursos gráficos ou multimídia. Bancos de dados orientados a objeto são mais apropriados para tal finalidade.

Um **DBMS orientado a objeto** armazena dados e procedimentos como se fossem objetos que podem ser automaticamente extraídos e compartilhados. Os sistemas de gestão de dados orientados a objeto (OODBMS — *object-oriented database management systems*) estão se tornando comuns porque podem ser usados para gerenciar os vários componentes multimídia ou *applets* Java utilizados em aplicativos da Web, que costumam integrar informações de uma variedade de fontes.

Embora os bancos de dados orientados a objeto possam armazenar tipos mais complexos de informação do que os DBMSs relacionais, são relativamente lentos quando se trata de processar grandes quantidades de transações. Sistemas híbridos de **DBMSs relacionais orientados a objeto** já estão disponíveis para prover recursos de DBMSs orientados a objeto e relacionais.

Como usar bancos de dados para melhorar o desempenho e a tomada de decisão na empresa

As empresas usam bancos de dados para monitorar transações básicas, como pagamento a fornecedores, processamento de pedidos, atendimento a clientes e pagamento a

Figura 5.10

Exemplo de uma consulta SQL

Vemos aqui os comandos SQL para uma consulta que selecionará fornecedores das peças 137 ou 150. Eles produzem uma lista com os mesmos resultados da Figura 5.8.

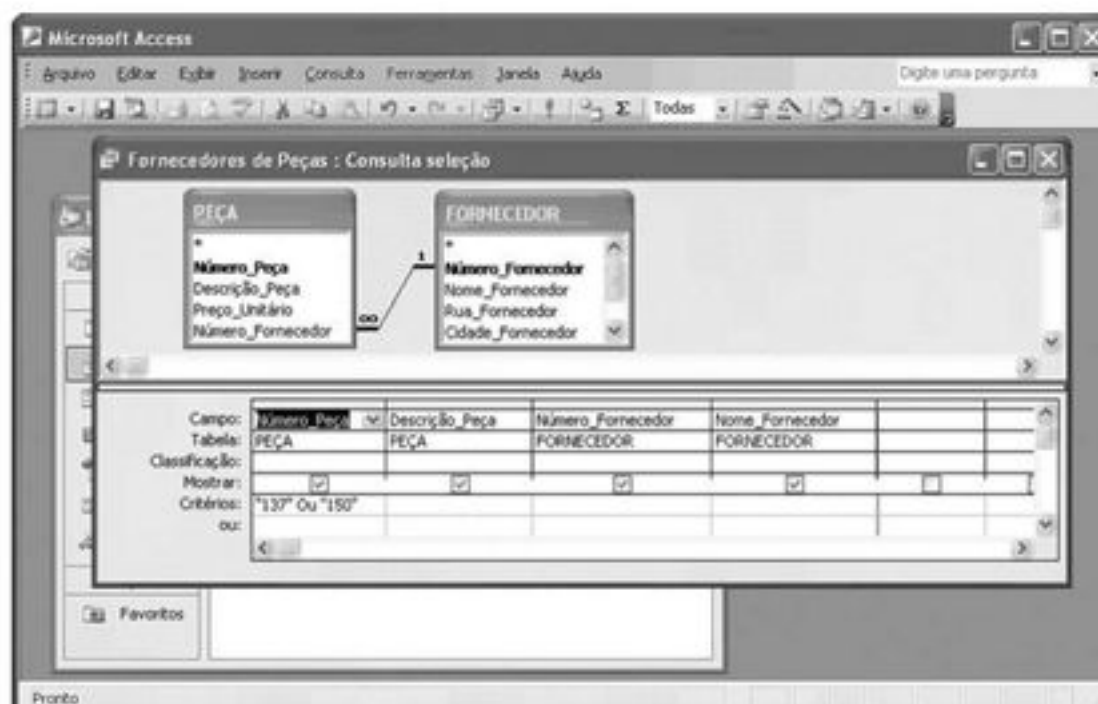


Figura 5.11

Uma consulta no Access

Vemos aqui como a consulta da Figura 5.10 seria elaborada com as ferramentas de formulação de consultas do Access. A tela mostra as tabelas, os campos e os critérios de seleção usados para a consulta.

funcionários, mas elas também precisam de bancos de dados para obter informações que as ajudem a administrar o negócio de maneira mais eficiente e, ao mesmo tempo, auxiliem gerentes e funcionários a tomar melhores decisões. Se uma empresa quiser saber qual produto é mais aceito pelo mercado, ou quais clientes são mais lucrativos, a resposta estará nos dados.

Por exemplo, ao analisar as informações das compras com cartão de crédito dos clientes, a Louise's Trattoria, uma cadeia de restaurantes de Los Angeles, percebeu que a qualidade era mais importante do que o preço para a maioria dos clientes, os quais tinham nível universitário e apreciavam vinhos refinados. Agindo a partir dessa informação, a cadeia introduziu pratos vegetarianos, estendeu as opções de frutos do mar e passou a oferecer vinhos mais caros, o que elevou as vendas em mais de dez por cento.

Em uma grande empresa, com grandes bancos de dados ou sistemas para funções separadas, como manufatura, vendas e contabilidade, são necessários recursos e ferramentas especiais para analisar vastas quantidades de dados e extraí-los de múltiplos sistemas. Entre esses recursos estão o *data warehousing* (armazenamento de dados), o *data mining* (mineração de dados) e ferramentas para acessar bancos de dados internos por meio da Web.

Data warehouses

O que você faria se quisesse informações concisas e confiáveis sobre operações correntes, tendências e mudanças relativas à empresa inteira? Se trabalhar em uma grande companhia, isso pode ser difícil, porque os dados geralmente são armazenados em sistemas separados, como vendas, manufatura ou contabilidade. Alguns dos dados de que precisa estarão no sistema de vendas, e outros no de manufatura. Para completar, muitos desses sistemas são legados com tecnologias de gestão de dados ultrapassadas, ou sistemas de arquivos nos quais a informação é de difícil acesso ao usuário.

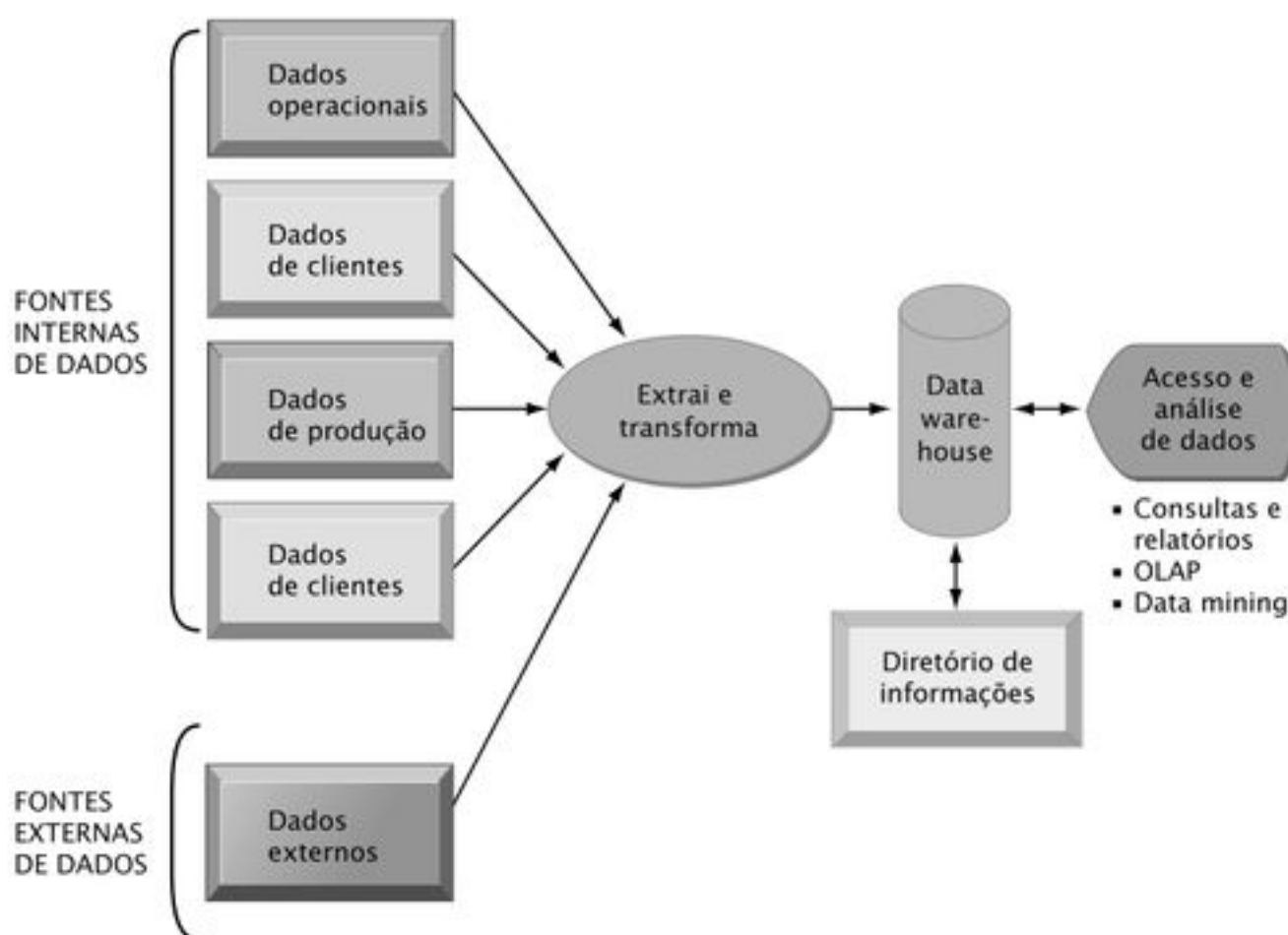
Talvez você gaste dias e dias localizando e reunindo os dados necessários, ou talvez seja obrigado a tomar sua decisão com base em informações incompletas. Se estiver pesquisando tendências, terá um problema adicional, porque a maioria das empresas deixa apenas seus dados atuais imediatamente disponíveis, sendo difícil encontrar dados sobre eventos passados. O *data warehousing* combate todos esses problemas.

O que é um data warehouse?

Data warehouse é um banco de dados que armazena dados correntes e históricos de potencial interesse para os tomadores de decisão de toda a empresa. Os dados originam-se de muitos sistemas operacionais centrais, como sistemas de vendas, contas de clientes e manufatura, podendo incluir ainda dados advindos de transações em sites. O Data warehouse consolida e padroniza as informações oriundas de diferentes bancos de dados operacionais, de modo que elas possam ser usadas por toda a empresa para análise gerencial e tomada de decisões.

A Figura 5.12 mostra como um data warehouse funciona. Os dados estão disponíveis a todos para acesso conforme a necessidade, mas não podem ser alterados. Um sistema de data warehouse também oferece uma gama de ferramentas *ad hoc* e padronizadas, como ferramentas para consulta, ferramentas analíticas e recursos para relatórios gráficos. Muitas empresas usam portais de intranet para disponibilizar as informações do data warehouse a todo o público interno.

A Seção Interativa sobre a Receita Federal norte-americana e seu data warehouse de compliance ilustra como data warehouses podem atuar como uma ferramenta poderosa para aumentar o poder de decisão e a eficiência operacional. Como os dados dos contribuintes estavam fragmentados entre muitos sistemas criados ao longo dos anos, a Receita Federal não conseguia formar uma imagem completa e abrangente dos contribuintes, tampouco analisar os dados com facilidade para identificar as pessoas propensas a mentir na declaração de sua renda. Um data warehouse permitiu integrar e centralizar os dados dos contribuintes e responder mais rapidamente às suas consultas. Ao ler o estudo de caso, tente identificar os problemas enfrentados pela Receita, as soluções alternativas disponíveis para a gerência e

**Figura 5.12****Componentes de um data warehouse**

Dados correntes e históricos são extraídos de sistemas operacionais internos à organização. Esses dados são combinados com dados de fontes externas e reorganizados em um banco central projetado para análise gerencial e produção de relatórios. O diretório de informações fornece aos usuários informações sobre os dados disponíveis no data warehouse.

as questões humanas, organizacionais e tecnológicas envolvidas que precisaram ser consideradas no desenvolvimento da solução.

Data marts

Empresas podem montar data warehouses organizacionais, nos quais um armazém central de dados atende à organização inteira, ou podem criar armazéns menores, descentralizados, denominados data marts. **Data mart** é um subconjunto de um data warehouse, no qual uma porção resumida ou altamente focalizada dos dados da organização é colocada em um banco separado destinado a uma população específica de usuários. Por exemplo, uma empresa pode desenvolver data marts de vendas e marketing para tratar informações de clientes. Um data mart em geral focaliza uma única área de interesse ou linha de negócios, de modo que pode ser montado com mais rapidez e a custo mais baixo do que um data warehouse de âmbito empresarial.

Inteligência empresarial, análise multidimensional de dados e data mining

Uma vez capturados e organizados em data warehouses e data marts, os dados ficam disponíveis para análises posteriores. O usuário conta com uma série de ferramentas para analisar essas informações e, assim, descobrir novos padrões, relacionamentos e insights úteis para orientar a tomada de decisão. Essas ferramentas que consolidam, analisam e acessam vastas quantidades de dados para ajudar os usuários a decidir melhores escolhas empresariais são, muitas vezes, chamadas de **inteligência empresarial (BI – business intelligence)**. Entre as principais ferramentas para inteligência empresarial, estão os softwares para consulta e relatórios de bancos de dados, as ferramentas para análise multidimensional de dados (processamento analítico on-line – OLAP) e o data mining.

Quando pensamos no conceito de *inteligência* aplicado aos seres humanos, normalmente presumimos a capacidade humana de combinar o conhecimento adquirido com novas informações e mudar comportamentos, a fim de executar determinada tarefa com sucesso ou adaptar-se a uma nova situação. De maneira análoga, a inteligência empresarial confere às empresas a capacidade de acumular informações; adquirir conhecimentos sobre clientes, concorrentes e operações internas; e mudar o comportamento de tomada de decisão a fim de alcançar maior lucratividade e outras metas corporativas.

SEÇÃO INTERATIVA: ORGANIZAÇÕES A Receita Federal descobre fraude nos impostos com data warehouse

A Receita Federal norte-americana é a agência responsável por coletar os impostos e definir as leis a eles relacionadas. Desde sua criação, na década de 1860, a Receita cresceu em ordens de magnitude equivalentes à da população dos Estados Unidos. Em 2008, a agência processou aproximadamente 140 milhões de declarações de impostos individuais que totalizaram uma receita de aproximadamente 2 trilhões de dólares. Não é surpresa alguma que uma falha em seus sistemas de informação poderia resultar em volumosa perda ao governo federal norte-americano. Felizmente para a Receita Federal — e, talvez, infelizmente para alguns inescrupulosos — a agência uniu-se à Sybase para implementar um data warehouse, conhecido como data warehouse de compliance (CDW — *compliance data warehouse*), que melhorou a eficiência e aumentou drasticamente o montante de dinheiro coletado pela Receita de seus contribuintes delinquentes.

A Receita precisava de um data warehouse para organizar suas informações acumuladas, com informações pessoais de contribuintes e suas declarações de rendimento. Os dados estavam armazenados em sistemas legados projetados para o processamento eficiente de formulários de declarações e organizados em muitos formatos diferentes, inclusive em bancos de dados hierárquicos de mainframes, bancos de dados relacionais em Oracle e arquivos sequenciais que não estavam no formato de banco de dados. Era quase impossível consultar e analisar os dados armazenados em bancos de dados hierárquicos e nos arquivos unidimensionais e não podiam ser facilmente combinados com os dados relacionais.

O CDW permite consultas altamente flexíveis em um dos maiores bancos de dados do mundo, com sete anos de informações de declarações de impostos individuais e empresariais. A cada ano, quatro terabytes de dados são inseridos no sistema. O banco de dados do data warehouse é relacional, com bilhões de linhas e mais de 200 colunas, todas com complexas ligações a agendas importantes e outros anexos. Quando chegam, os dados são reorganizados na estrutura relacional com base em definições e formatos padrão. Os pesquisadores da Receita agora conseguem pesquisar e analisar centenas de milhões, ou mesmo bilhões, de registros de uma vez utilizando fonte centralizada de dados precisos e consistentes, em vez de terem de agrupar informações de fontes inconsistentes.

A implantação do CDW aprimorou muitíssimo a capacidade de a Receita gerenciar e utilizar os dados coletados. Como resultado, o data warehouse permitiu que a agência recuperasse muitos bilhões de dólares em declarações de rendimento que ficavam perdidas no antigo sistema. Em 2006, por exemplo, a Receita coletou 59,2 bilhões de dólares em receita adicional através de 1,4 milhão de auditorias de contribuintes investigados por declarações que não relatavam todos os rendimentos recebidos.

A capacidade do CDW aumentou de três terabytes desde sua criação, no final da década de 1990, para 150 terabytes de dados atualmente. Ele permite que os usuários consultem os dados através de uma série de ferramentas. Inicialmente, o CDW contava com Sybase Adaptive Server IQ (software de dados relacional para data warehouses atualmente denominado Sybase IQ), Sybase PowerBuilder (ferramenta para desenvolvimento de aplicações para que usuários listassem e acessassem o conteúdo do banco de dados), Sybase Open Client (interface entre os sistemas do cliente e os servidores Sybase), Open Database Connectivity (interfaces de programação da aplicação), servidores Dual Sun Enterprise 6000 rodando Solaris 2.6 (versão da Sun para o UNIX), e disc array EMC. Esperava-se que o recurso mais importante do data warehouse fosse suficientemente grande para acomodar múltiplos terabytes de dados, mas também acessível de modo a permitir consultas a seus dados a partir de diferentes ferramentas. Os componentes selecionados pela Receita permitiram que o CDW funcionasse exatamente dessa forma.

A implantação do CDW não se deu sem desafios. Um dos maiores deles foi descobrir que a conversão dos dados legados para o novo sistema não era um processo uniforme. Como as leis fiscais mudaram muitas vezes ao longo dos anos, a estrutura dos dados da Receita não era consistente de ano para ano; o que tornou a integração dos dados um processo complicado. Além disso, o montante de dados que o CDW estava disposto a gerenciar era muito maior do que qualquer outro que a Receita já houvesse gerenciado. Convencer a realizar uma atualização tão abrangente quanto a implantação de um data warehouse também não era fácil, já que as agências governamentais costumam ser avessas aos riscos e resistem a mudanças desse tipo. Os data warehouses também tendem a demandar esforço extensivo e dinheiro para se manterem atualizados.

Apesar dos obstáculos, a implantação foi um grande sucesso. A Receita relata que alcançou retorno de 200 por 1 na taxa de investimento pouco depois da implantação do CDW, que custou somente 2 milhões de dólares para ser concluído. A maior parte das economias do CDW veio da velocidade e da facilidade com as quais o sistema detectava erros nas declarações de rendimento. Utilizando o data warehouse, os analistas conseguem determinar padrões em grupos de pessoas propensas a mentir em suas declarações, tais como casais divorciados em que ambos declaram os filhos em seus formulários do mesmo ano, pessoas que abusam das deduções ou dos paraísos fiscais para pequenas empresas, ou recém-graduados sobrecarregados com empréstimos estudantis que podem acabar pagando menos impostos. Houve redução de seis a oito meses para somente algumas horas do tempo necessário à identificação de erros nas reclamações e para a análise de dados.

Mais recentemente, a Receita atualizou a maneira de transportar as informações para o data warehouse central. Nos estágios iniciais, de desenvolvimento, a agência transportava os dados utilizando fitas magnéticas que comportavam somente dois gigabytes cada. Em 2006, a Receita substituiu as fitas por dispositivos de armazenamento de dois terabytes conectados à rede, com tamanho semelhante ao das fitas, mas que armazenam o volume de dados equivalente a 1.500 fitas. Além disso, os dispositivos de armazenamento são criptografados, garantindo que os dados fiquem seguros durante o transporte; anteriormente as fitas não eram seguras e deixavam as informações dos contribuintes desprotegidas durante o trânsito. Estima-se que essa modificação represente uma economia de milhões de dólares para a agência ao longo do período de cinco anos.

O número de auditorias realizadas pela Receita sugere que o CDW está funcionando bem, resultando

em mais auditorias dos fraudadores e menos auditorias dos contribuintes honestos. As chances de sofrer auditoria aumentaram de uma em 140 em 2006 para uma em 377 em 2000. Em 2006, os contribuintes que receberam 1 milhão de dólares ou mais anualmente tiveram uma chance em 11 de ser auditados. Em 2003, as chances eram de uma em 20 para a mesma faixa de rendimentos. Contudo, a Receita conseguiu reduzir o número de auditorias realizadas nos dados de contribuintes inocentes e, portanto, o crescente número de auditorias afetou primeiramente aqueles que estão em dívida.

Fontes: Eric Lai, "Been audited lately? Blame the IRS's massive, superfast data warehouse". *Computerworld*, 22 mar. 2008; www.irs.gov, acesso em 31 jul. 2009; Sybase, "Internal Revenue Service". www.sybase.com, acesso em 25 mar. 2008; e "Agency award: The IRS' update of the compliance data warehouse makes analysis less taxing". *Government Computer News*, 8 out. 2007.

PERGUNTAS SOBRE O ESTUDO DE CASO

1. Por que era tão difícil para a Receita Federal analisar os dados coletados dos contribuintes?
2. Quais desafios a Receita Federal encontrou durante a implementação de seu data warehouse de compliance? Quais questões gerenciais, organizacionais e tecnológicas precisaram ser consideradas?
3. De que maneira o CDW aprimorou a tomada de decisão e as operações na Receita Federal? Existem benefícios aos contribuintes?
4. Você acha que os data warehouses poderiam ser úteis em outras áreas do setor federal? Quais? Explique.

Por exemplo, a Harrah's Entertainment, a segunda maior empresa do setor de jogos, analisa continuamente dados sobre os clientes, coletados enquanto as pessoas jogam nos caça-níqueis da empresa ou frequentam cassinos e hotéis da rede. O departamento de marketing da Harrah's utiliza essas informações para desenvolver um detalhado perfil do jogador, com base no valor ao longo do tempo de um cliente particular à empresa. Essa informação orienta as decisões administrativas sobre como cultivar os clientes mais lucrativos, estimulá-los a gastar mais e atrair mais clientes com alto potencial de geração de receita. A inteligência empresarial elevou tanto os lucros da Harrah's que se tornou peça central de sua estratégia de negócios.

A Figura 5.13 ilustra como funciona a inteligência empresarial. Os bancos de dados operacionais da empresa monitoram as transações geradas no dia a dia dos negócios e, com suas informações, alimentam o data warehouse. Os gerentes usam ferramentas de inteligên-

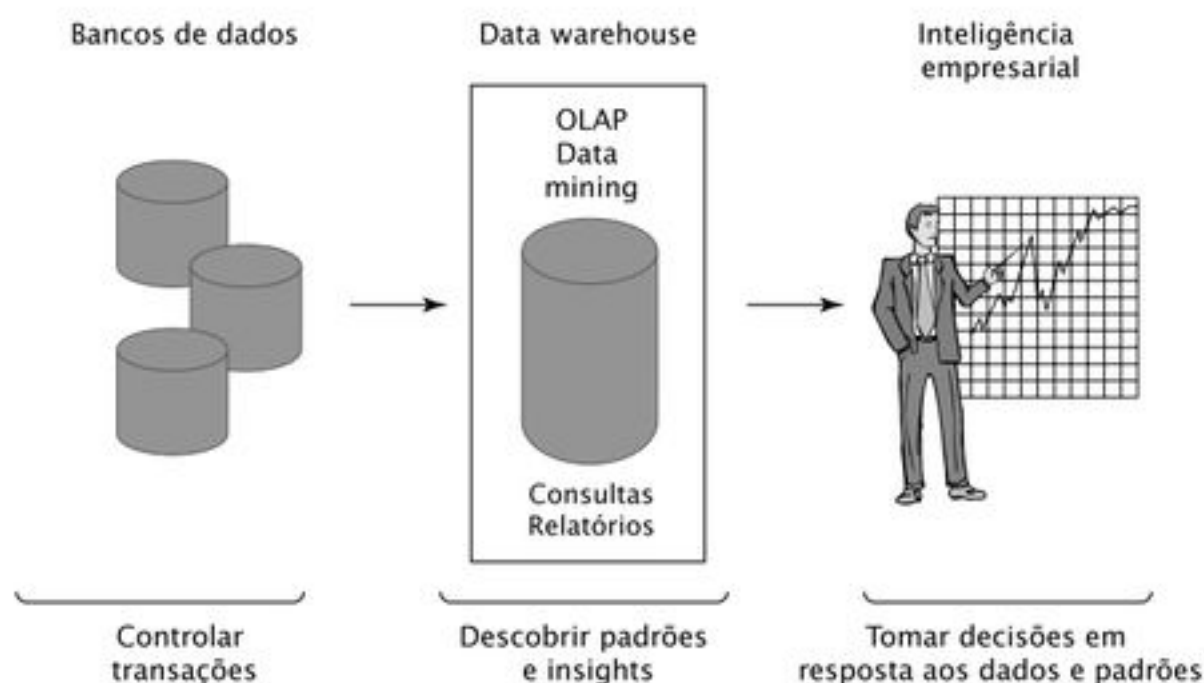


Figura 5.13

Inteligência empresarial

Uma série de ferramentas analíticas trabalha com os dados armazenados nos bancos de dados, encontrando padrões e insights que ajudam gerentes e funcionários a tomar melhores decisões e, assim, aprimorar o desempenho organizacional.

cia empresarial para encontrar sentidos e padrões nos dados. Podem, então, agir segundo o que aprenderam com a análise, tomando decisões mais bem fundamentadas e inteligentes.

Esta seção apresenta as mais importantes ferramentas e tecnologias de inteligência empresarial. Oferecemos mais detalhes sobre aplicativos de inteligência empresarial no Capítulo 10, quando abordamos a tomada de decisão.

Processamento analítico on-line (OLAP – on-line-analytical processing)

Imagine que sua empresa venda quatro produtos diferentes — porcas, pinos, arruelas e parafusos —, nas regiões leste, oeste e central. Se quisesse fazer uma pergunta bem direta, como “Quantas arruelas foram vendidas durante o último trimestre?”, poderia facilmente achar a resposta consultando seu banco de dados de vendas. Mas e se quisesse saber quantas arruelas foram vendidas em cada região e também comparar os resultados reais com a projeção de vendas?

Para obter essa resposta, é necessário um **processamento analítico on-line (OLAP)**. O OLAP permite a análise multidimensional de dados, de forma que os usuários vejam os mesmos dados de diferentes maneiras, pois usa múltiplas dimensões. Cada aspecto da informação — produto, preço, custo, região ou período de tempo — representa uma dimensão diferente. Dessa forma, um gerente de produto poderia utilizar uma ferramenta de análise multidimensional de dados para saber quantas arruelas foram vendidas na região leste em junho, comparar essas informações com as do mês anterior e também de junho do ano anterior e ainda confrontá-las com a previsão de vendas. O OLAP permite que os usuários obtenham respostas on-line sobre questões específicas como essas em velocidade razoável, mesmo quando os dados estão armazenados em bancos gigantescos, como números de vendas de vários anos.

A Figura 5.14 mostra um modelo multidimensional que poderia ser criado para representar produtos, regiões, volume real de vendas e projeção de vendas. Uma matriz de vendas reais pode ser empilhada sobre uma matriz de projeção de vendas formando um cubo com seis faces.

Girando o cubo 90 graus em uma direção, a face que aparece mostrará vendas reais e projetadas por produto. Girando novamente o cubo 90 graus, podem-se ver vendas reais e projetadas por região. A 180 graus a partir da visão original podem-se visualizar vendas projetadas e produtos por região. Os cubos podem ser aninhados dentro de outros cubos para construir visões complexas dos dados. Uma empresa pode usar um banco de dados multidimensional especializado, ou uma ferramenta que crie visões multidimensionais em bancos de dados relacionais.

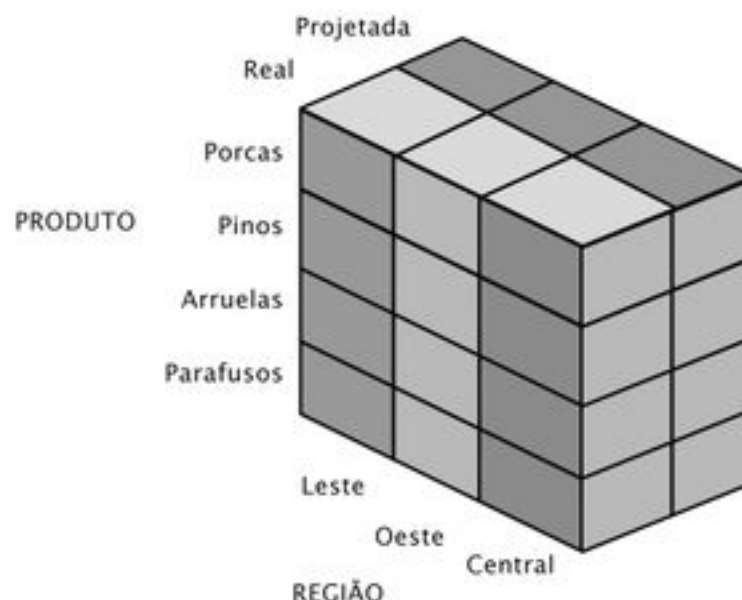
Data mining

Consultas tradicionais a bancos de dados respondem a perguntas como: “quantas unidades do produto número 403 foram expedidas em fevereiro de 2010?”. O OLAP, ou a análise

Figura 5.14

Modelo multidimensional de dados

A visão mostrada é produto x região. Girando o cubo 90 graus, a face que aparecer mostrará produto x vendas projetadas e reais. Girando novamente o cubo 90 graus, pode-se ver região x vendas reais e projetadas. Outras visões também são possíveis.



multidimensional, dá apoio a requisições de informação muito mais complexas, tais como: “compare as vendas do produto 403 com o planejado, por trimestre e por região de vendas, nos dois últimos anos”. Para utilizar o OLAP e a análise de dados orientada por consulta, os usuários precisam ter uma boa ideia da informação que estão procurando.

O **data mining** é mais orientado por descoberta. O data mining fornece percepções dos dados corporativos que não podem ser obtidas com o OLAP, descobrindo padrões e relacionamentos ocultos em grandes bancos de dados e inferindo regras a partir deles para prever comportamentos futuros. Esses modelos e regras podem então ser utilizados para guiar o processo de decisão e prever o efeito dessas escolhas. Os tipos de informação obtidos com o data mining incluem associações, sequências, classificações, aglomerações e prognósticos.

- *Associações* são ocorrências ligadas a um único evento. Por exemplo, um estudo de modelos de compra em supermercados pode revelar que, na compra de salgadinhos de milho, compra-se também um refrigerante tipo cola em 65 por cento das vezes; mas, quando há uma promoção, o refrigerante é comprado em 85 por cento das vezes. Com essa informação, os gerentes podem tomar decisões mais acertadas, pois aprenderam a respeito da rentabilidade de uma promoção.
- Na *sequência*, os eventos estão ligados ao longo do tempo. Pode-se descobrir, por exemplo, que, quando se compra uma casa, em 65 por cento das vezes se adquire também uma nova geladeira no período de duas semanas; e que, em 45 por cento das vezes, um fogão também é comprado um mês após a compra da residência.
- A *classificação* reconhece modelos que descrevem o grupo ao qual o item pertence, por meio do exame dos itens já classificados e pela inferência de um conjunto de regras. Empresas de operadoras de cartões de crédito e companhias telefônicas preocupam-se com a perda de clientes regulares. A classificação pode ajudar a descobrir as características de clientes que provavelmente virão a abandoná-las e oferecer um modelo para ajudar os gerentes a prever quem são, de modo que se elabore antecipadamente campanhas especiais para reter esses clientes.
- A *aglomeração (clustering)* funciona de maneira semelhante à classificação quando ainda não foram definidos grupos. Uma ferramenta de data mining descobrirá diferentes agrupamentos dentro da massa de dados; por exemplo, ao encontrar grupos de afinidade para cartões bancários ou ao dividir o banco de dados em categorias de clientes com base na demografia e em investimentos pessoais.
- Embora todas essas aplicações envolvam previsões, os *prognósticos* as utilizam de modo diferente. Partem de uma série de valores existentes para prever quais serão os outros valores. Por exemplo, um prognóstico pode descobrir padrões nos dados que ajudam os gerentes a estimar o valor futuro de variáveis como números de vendas.

Esses sistemas realizam uma análise de alto nível quanto a padrões ou tendências, mas também podem esmiuçar os dados para revelar mais detalhes, se necessário. Existem aplicações de data mining para todas as áreas funcionais da empresa, bem como para o trabalho científico ou governamental. É comum usar o data mining para analisar detalhadamente padrões em dados sobre consumidores e, a partir disso, montar campanhas de marketing um-a-um ou identificar clientes lucrativos.

A Virgin Mobile Austrália usa um data warehouse e o data mining para aumentar a fidelidade dos clientes e lançar novos serviços. A empresa criou um data warehouse que consolidou os dados do sistema integrado, do sistema de gestão do relacionamento com o cliente e dos sistemas de cobrança em um gigantesco banco de dados. Graças ao data mining, a administração pôde determinar o perfil demográfico dos novos clientes e relacioná-lo aos tipos de aparelho telefônico que compravam, assim como identificar o desempenho de cada loja e das campanhas no ponto de venda, as reações dos clientes a novos produtos e serviços, as taxas de cancelamento com o cliente e a receita gerada segundo cada cliente.

A **análise preditiva** usa técnicas de data mining, dados históricos e suposições sobre condições futuras para prever resultados de eventos, tais como a probabilidade de um cliente responder a uma oferta ou comprar um produto específico. A divisão norte-americana da The Body Shop International S.A., por exemplo, fazia análise preditiva em seu

banco de dados, examinando informações sobre clientes de catálogo, da Web e do varejo para identificar aqueles mais inclinados a fazer compras por catálogo. Essas informações ajudaram a empresa a montar uma mala direta mais precisa e direcionada para seus catálogos, aumentando a taxa de resposta e as receitas desse canal.

O data mining é uma ferramenta poderosa e lucrativa, mas impõe desafios em relação à proteção da privacidade. A tecnologia de data mining pode combinar informações oriundas de diversas fontes para criar uma 'imagem digital' detalhada de cada um — nossa renda, nossos padrões de locomoção, nossos *hobbies*, nossa família, nossos interesses políticos. No Capítulo 12, consideramos se as empresas devem ou não ter permissão para coletar informações tão detalhadas sobre os indivíduos.

Text mining (mineração de texto) e *Web mining* (mineração na Web)

As ferramentas de inteligência empresarial lidam primeiramente com as informações estruturadas em bancos de dados e arquivos. Acredita-se, entretanto, que dados não estruturados, em sua maioria, na forma de arquivos de texto, representam mais de 80 por cento das informações úteis de uma empresa. E-mails, memorandos, transcrições de chamadas, respostas para pesquisas, casos judiciais, descrições de patentes e relatórios de serviços são valiosos na identificação de padrões e tendências capazes de ajudar os empregados a tomar decisões de negócios mais acertadas. Ferramentas para **mineração de textos** (*text mining*) agora estão disponíveis para auxiliar as empresas na análise desses dados. Essas ferramentas são capazes de extrair elementos-chave de grandes conjuntos de dados não estruturados, descobrir padrões e relacionamentos, e resumir informações. As empresas podem recorrer à mineração de textos para analisar transcrições de ligações às centrais de atendimento ao cliente a fim de identificar os principais problemas relacionados a serviços e reparos.

A empresa Air Products and Chemicals, em Allentown, Pensilvânia, utiliza a mineração de textos para ajudar na identificação de documentos que requerem procedimentos especiais de retenção para atender às exigências da Lei Sarbanes-Oxley. A empresa possui mais de nove terabytes de dados não estruturados (sem incluir e-mails). O software SmartDiscovery, da Inxight Software, classifica e organiza essas informações de modo que a empresa possa aplicar regras de negócios a uma categoria de documentos, em vez de aplicá-las a documentos individuais. Se for descoberto que um documento trata de operações incluídas na Lei Sarbanes-Oxley, a empresa vai se certificar de que o documento atende aos requisitos de retenção de dados expressos pela Lei.

A Web é outra fonte rica em informações valiosas, e algumas podem ser pesquisadas para encontrar padrões, tendências e observações sobre o comportamento do consumidor. A descoberta e análise de padrões e informações úteis na World Wide Web é chamado de **mineração na Web** (*web mining*). As empresas podem recorrer à mineração na Web para ajudar-lhes a compreender o comportamento do consumidor, avaliar a eficiência de um site em particular ou quantificar o sucesso de uma campanha de marketing. Os comerciantes usam o Google Trends e o Google Insight, por exemplo, para serviços de busca, que rastreiam a popularidade de diversas palavras e frases usadas nas consultas feitas pelo Google para descobrir no que as pessoas estão interessadas e o que elas desejam comprar.

A mineração na Web busca padrões em dados através da mineração de conteúdo, de estrutura e de uso. A mineração de conteúdo é o processo de extração de conhecimento a partir do conteúdo de páginas da Web, que pode incluir dados nos formatos texto, imagem, áudio e vídeo. A mineração de estrutura examina os dados relacionados à estrutura de um site em particular. Por exemplo, links apontando para um documento indicam a popularidade daquele documento, enquanto links saindo de um documento representam a riqueza, ou a variedade, de tópicos abordados no documento. A mineração de uso analisa os dados de interação do usuário registrados por um servidor sempre que recebidas solicitações de recursos de um site. Os dados sobre o uso registram o comportamento do usuário quando ele consulta ou realiza transações na internet e grava esses dados em um *log* no servidor. A análise de tais dados permite que as empresas determinem o valor de determinados clientes, as estratégias cruzadas de marketing entre os produtos e a eficiência de campanhas promocionais.

Bancos de dados e a Web

Muitas empresas estão usando a Web para disponibilizar algumas informações de seus bancos de dados internos a clientes e parceiros de negócios. Clientes potenciais podem usar o site de uma empresa para ver o catálogo de produtos ou fazer um pedido. A empresa pode então usar a Web para verificar a disponibilidade daquele produto no estoque de seu fornecedor. O fornecedor, por sua vez, pode acionar seus próprios fornecedores, assim como as empresas de entrega necessárias para enviar os produtos no prazo.

Essas ações envolvem o acesso e (no caso do pedido) a atualização de bancos de dados corporativos por meio da Web. Suponha, por exemplo, que um cliente com um navegador da Web queira pesquisar preços no banco de dados on-line de um varejista. A Figura 5.15 mostra como esse cliente poderia acessar o banco de dados interno do varejista pela rede. Usando um software de navegação da Web no seu PC, o cliente acessaria o site do varejista pela Internet. O software de navegação da Web do usuário solicitaria informações do banco de dados da organização, usando comandos HTML para se comunicar com o servidor da Web.

Como muitos bancos de dados *back-end* não conseguem interpretar comandos escritos em HTML, o servidor da Web repassaria essas solicitações para o software que traduz comandos HTML em SQL, de maneira que elas pudessem ser processadas pelo DBMS que trabalha com o banco de dados. Em um ambiente cliente/servidor, o DBMS muitas vezes reside em um computador dedicado chamado **servidor de banco de dados**. O DBMS recebe os pedidos SQL e fornece os dados requisitados. O *middleware* transfere então as informações do banco de dados interno da organização de volta para o servidor da Web, exibindo-as na forma de uma página da rede.

A Figura 5.15 mostra que o software que trabalha entre o servidor Web e o DBMS poderia ser um servidor de aplicativo rodando em seu próprio computador dedicado (ver Capítulo 4). O software servidor de aplicativo administra todas as operações de aplicativos, incluindo o processamento de transações e o acesso aos dados, entre computadores munidos de navegador e as aplicações empresariais de suporte ou o banco de dados da empresa. O servidor de aplicativo pega as requisições do servidor da Web, executa a lógica empresarial das transações tendo como base essas requisições e provê conexão aos sistemas de suporte ou bancos de dados da organização. De modo alternativo, essas operações podem ser administradas por um programa de software customizado ou um script CGI, programa compacto que usa a especificação *Common Gateway Interface (CGI)* para processar dados em um servidor da Web.

Há uma série de vantagens em utilizar a rede para acessar os bancos de dados internos de uma organização. Em primeiro lugar, todo mundo sabe usar um software de navegação da Web, e os funcionários precisam de muito menos treinamento do que se usassem ferramentas de consulta proprietárias. Em segundo lugar, a interface requer pouca ou nenhuma mudança no banco de dados interno. As empresas alavancam seus investimentos em sistemas mais antigos porque custa muito menos adicionar uma interface da Web à frente de um sistema legado do que reprojeter e reconstruir o sistema para melhorar o acesso do usuário. Por essa razão, a maioria das grandes empresas no ranking da *Fortune 500* usa bancos de dados legados de *back-end* rodando em mainframes ligados a um software de *front-end*, que disponibiliza as informações na forma de uma página da Web aos usuários que as solicitam.

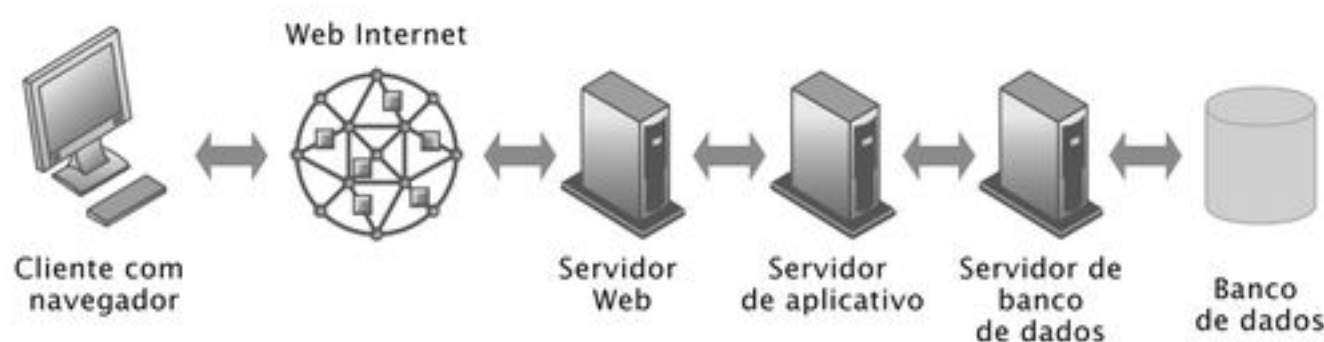


Figura 5.15

Ligando bancos de dados internos à Web

Usando seus PCs e software de navegação da Web, os usuários acessam o banco de dados interno de uma organização por meio da Internet.

Acessar bancos de dados corporativos pela Web está criando novos rendimentos e novas oportunidades, em alguns casos até mesmo mudando a maneira como os negócios são feitos. A ThomasNet.com, por exemplo, oferece um diretório com informações atualizadas referentes a mais de 650 mil fornecedores de produtos industriais, tais como produtos químicos, metais, plásticos, produtos de borracha e equipamentos automotivos. Antes chamada Thomas Register, a empresa costumava enviar imensos catálogos impressos com tais informações. Atualmente, disponibiliza os dados on-line em seu site, tendo se tornado uma organização mais enxuta.

Outras empresas criam negócios totalmente novos com base no acesso a grandes bancos de dados pela Web. O site MySpace, por exemplo, ajuda os usuários a conhecerem outras pessoas ou se manterem conectados uns aos outros. MySpace oferece músicas, comédia, vídeos e “perfis” com informações fornecidas por mais de 175 milhões de usuários sobre idade, cidade natal, preferências para encontros, estado civil e interesses. O site mantém um grande banco de dados para abrigar e gerenciar todo esse conteúdo.

Gestão dos recursos de dados

Configurar um banco de dados é somente o início. Para certificar-se de que os dados para o seu negócio permanecem precisos, confiáveis e prontamente disponíveis para os que deles necessitam, sua empresa necessitará de políticas e procedimentos especiais sobre gestão de dados.

Estabelecendo uma política de informação

Todas as empresas, grandes ou pequenas, precisam de uma política de informação. Os dados de sua organização são um recurso importante, por isso você não vai querer que os outros façam o que quiserem com eles. É necessário estabelecer regras sobre como os dados serão organizados e armazenados, e quem terá permissão para vê-los ou alterá-los.

Uma **política de informação** especifica as regras para compartilhar, disseminar, adquirir, padronizar, classificar e inventariar a informação. A política de informação elabora procedimentos e responsabilidades específicas, determinando quais usuários e unidades organizacionais compartilham a informação, para onde ela pode ser distribuída e quem é responsável por sua atualização e manutenção. Por exemplo, uma política de informação típica especificaria que apenas membros selecionados do departamento de recursos humanos e da folha de pagamento teriam o direito de alterar e visualizar informações delicadas sobre os funcionários, tais como número da Previdência Social ou salário, e esses departamentos seriam responsáveis por assegurar a precisão de tais dados.

Caso trabalhe em uma empresa pequena, a política de informação talvez seja estabelecida e implementada pelos proprietários ou gerentes. Em uma grande organização, o planejamento e a gestão da informação como recurso corporativo muitas vezes requerem uma função formal de administração de dados. A **administração de dados** é responsável pelas políticas e procedimentos específicos pelos quais as informações podem ser gerenciadas como recurso organizacional. Essas responsabilidades incluem desenvolvimento da política de informação, planejamento de dados, supervisão do projeto lógico do banco de dados e do desenvolvimento do dicionário de dados, e monitoração de como os especialistas em sistema de informação e grupos de usuários finais utilizam essas informações.

Uma grande organização também costuma ter um grupo de gestão e projeto de bancos de dados, dentro da divisão corporativa de sistemas de informação responsável por definir e organizar a estrutura e o conteúdo do banco de dados, e também por sua manutenção. Em estreita colaboração com os usuários, o grupo de projeto determina o banco de dados físico, as relações lógicas entre elementos e as regras de acesso e procedimentos. As funções que o grupo desempenha são chamadas de **gestão de banco de dados**.

Assegurando a qualidade de dados

Um banco de dados e uma política de informação bem projetados já são meio caminho andado para que a empresa tenha a informação de que precisa. No entanto, outras medidas devem ser tomadas para assegurar que os dados nos bancos de dados organizacionais sejam necessários e permaneçam confiáveis.

O que aconteceria se o número de telefone ou o saldo bancário de um cliente estivesse incorreto? Qual seria o impacto se um produto vendido estivesse com o preço errado no banco de dados? Informações incorretas, desatualizadas ou inconsistentes com outras fontes de informação criam sérios problemas operacionais e financeiros para as empresas. Quando dados errôneos passam despercebidos, muitas vezes levam a decisões incorretas, *recall* de produtos e até prejuízos financeiros.

Os consultores do Gartner Group relataram que mais de 25 por cento dos dados críticos nos bancos de dados das mil maiores empresas do ranking da *Fortune* estão incorretos ou incompletos; isso inclui erros em descrições e códigos de produtos, descrições de estoque, dados financeiros e informações sobre fornecedores e funcionários. Um estudo da Syrius Decision intitulado '*The impact of bad data on demand creation*' (em português, 'O impacto dos dados errôneos na criação de demanda') descobriu que de 10 a 25 por cento dos dados de clientes reais e potenciais contêm erros críticos. A correção desses erros em suas fontes e a adesão às melhores práticas para promoção da qualidade dos dados aumentou a produtividade dos processos de vendas e gerou aumento de 66 por cento na receita (Lager, 2009; Gage e McCormick, 2005).

Alguns desses problemas com qualidade dos dados são causados por dados inconsistentes e redundantes produzidos por múltiplos sistemas. Por exemplo, tanto o sistema de pedidos de vendas quanto o sistema de gestão de estoques podem manter informações sobre os produtos da organização. No entanto, o sistema de pedidos de vendas pode usar o termo *Número de Item* e o sistema de estoques pode chamar o mesmo atributo de *Número de Produto*. Os sistemas de vendas, estoque e manufatura de uma loja de roupas podem usar diferentes códigos para representar os valores de um mesmo atributo. Um deles pode designar o tamanho das roupas como 'extragrande', enquanto o outro pode usar o código 'XL' para a mesma finalidade. Durante o processo de projeto do banco de dados, aqueles que descrevem as entidades, como um cliente, um produto ou um pedido, devem ser nomeados e definidos de maneira consistente para todas as áreas da empresa que usarão o banco de dados.

Se o banco de dados for adequadamente projetado e forem estabelecidos padrões de dados para toda a empresa, a ocorrência de duplicações ou inconsistências será mínima. Contudo, a maioria dos problemas de qualidade de dados, como nomes grafados incorretamente, números trocados ou códigos faltantes ou incorretos, deriva de erros durante a entrada de dados. A incidência de tais erros está crescendo à medida que as empresas transferem parte de seu negócio para a Web e permitem que clientes e fornecedores insiram dados em seus sites, atualizando diretamente os sistemas internos.

Pense em todas as vezes que recebeu a mesma mala direta publicitária no mesmo dia. Muito provavelmente, isso ocorreu porque seu nome foi armazenado várias vezes no mesmo banco de dados. Seu nome pode ter sido grafado incorretamente, ou você pode ter usado a inicial do nome do meio em determinada ocasião e não em outra, ou a informação pode ter sido inicialmente inserida em um formulário impresso e escaneada inadequadamente para o sistema. Devido a essas inconsistências, o banco de dados o trata como se você fosse várias pessoas! Muitas vezes nós, como pessoas físicas, recebemos e-mails redundantes endereçados a Laudon, Lavdon, Lauden ou Landon.

Antes de implantar um novo banco de dados, as organizações precisam identificar e corrigir seus dados incorretos, além de estabelecer rotinas mais avançadas para editá-los quando o banco de dados estiver em operação. A análise da qualidade dos dados normalmente começa com uma **auditoria de qualidade de dados**, que consiste em levantamento estruturado da precisão e do nível de integridade dos dados em um sistema de informação. Auditorias desse gênero podem ser executadas com um levantamento completo dos

arquivos de dados, de amostras desses arquivos ou da percepção dos usuários finais quanto à qualidade dos dados.

O **data cleansing** (limpeza e padronização) consiste em atividades para detectar e corrigir, dentro do banco de dados, informações incorretas, incompletas, formatadas inadequadamente ou redundantes. O data cleansing não apenas corrige os dados, mas também reforça a consistência entre diferentes conjuntos de dados oriundos de sistemas de informação independentes. Softwares especializados em data cleansing, disponíveis comercialmente, podem automaticamente pesquisar arquivos de dados, corrigir erros nos dados e integrá-los em formato legível para toda a empresa.

Os problemas com qualidade de dados não são somente empresariais. Eles também representam sérios problemas aos indivíduos, afetando sua condição financeira e até mesmo seu emprego. A Seção Interativa sobre pessoas descreve alguns desses impactos ao detalhar problemas de qualidade nas empresas que coletam e relatam dados de crédito do consumidor. Ao ler o caso, procure as questões humanas, organizacionais e tecnológicas por trás do problema e avalie se as soluções existentes são adequadas.

SEÇÃO INTERATIVA: PESSOAS Erros no Bureau de Crédito — Problemas de gente grande

Você encontrou o carro de seus sonhos, possui um bom emprego e dinheiro suficiente para uma entrada e tudo de que precisa é um empréstimo no valor de 14 mil dólares. Tem poucas contas nos cartões de crédito, que paga fielmente todo mês. Quando se candidata ao empréstimo, entretanto, seu pedido é negado. Indagado o motivo, informam que possui um empréstimo em atraso junto a um banco do qual nunca ouviu falar. Você acabou de se tornar uma das milhões de pessoas vítimas de dados imprecisos ou desatualizados nos sistemas de informação das agências de crédito.

A maior parte dos dados sobre o histórico de crédito dos consumidores norte-americanos é coleta e mantida por três agências de crédito: Experian, Equifax e TransUnion. Essas empresas coletam dados de diversas fontes para criar um dossiê detalhado sobre os hábitos de empréstimos e pagamentos dos indivíduos. Essas informações ajudam os credores a avaliar o mérito de crédito de uma pessoa, a possibilidade de pagar o empréstimo, e podem afetar as taxas de juros e outros termos do empréstimo, a começar pela possibilidade de concedê-lo. As informações podem também afetar as chances de encontrar um emprego, já que, cada vez mais, os empregadores verificam os relatórios de crédito antes de contratar novos empregados.

As agências de crédito norte-americanas coletam informações pessoais e dados financeiros de diversas fontes com credores, serviços públicos, agências de recebimento de débitos e tribunais. Esses dados são agrupados e armazenados em grandes bancos de dados mantidos pelas agências de crédito que, por sua vez, vendem as informações para outras empresas para uso na avaliação de crédito.

As agências de crédito afirmam saber quais cartões de crédito estão na carteira de cada consumidor, quanto da hipoteca é devido e se a conta de luz é paga em dia. Mas se a informação errada for parar em seus sistemas, seja por roubo de identidade ou por erros nos dados

informados pelos credores, esteja atento! Desfazer a confusão pode ser quase impossível.

As agências compreendem a importância de fornecer informações precisas tanto para credores quanto para consumidores. Entretanto, também sabem que seus próprios sistemas são responsáveis pela maioria dos erros nos relatórios de crédito. Alguns deles ocorrem por conta dos procedimentos de compatibilização entre o empréstimo e o relatório individual de crédito.

O alto volume de informações transmitidas pelos credores às agências de crédito também aumenta a probabilidade de erros. A Experian, por exemplo, atualiza 30 milhões de relatórios de crédito por dia e aproximadamente 2 bilhões de relatórios ao mês. Ela compara a informação pessoal identificada em uma solicitação ou conta de crédito com a informação pessoal identificada em arquivo de crédito de cliente. A identificação de informações pessoais inclui itens como primeiro e último nomes, inicial do nome do meio, endereço atual completo e código postal, endereço anterior completo e código postal, e número do seguro social. As novas informações de crédito vão para o arquivo de informações de crédito do consumidor que melhor combina durante a comparação.

As agências de crédito dificilmente obtêm uma comparação ótima entre todos os campos e, portanto, precisam definir o quanto de variação permitir nas equivalências. Um cliente pode fornecer informações incompletas ou imprecisas em uma solicitação de crédito. Um credor pode enviar informações incompletas ou imprecisas às agências. Se a pessoa errada apresentar uma equivalência melhor do que qualquer outra, os dados vão para a conta errada, infelizmente.

Talvez o consumidor não tenha preenchido de forma clara a solicitação de empréstimo. Variações de nome em diferentes contas também podem resultar em combinações não tão perfeitas. Considere o nome Edward Jeffrey Johnson. Uma conta pode relatar Edward Johnson, outra pode informar Ed Johnson simples-

mente, enquanto uma terceira pode relatar Edward J. Johnson. Imagine que os dois últimos dígitos do CPF de Edward sejam trocados — mais chances de erro.

Se o nome ou o número do CPF de outra pessoa combinar parcialmente com os dados de seu arquivo, o computador pode anexar os dados dessa pessoa ao seu registro. Seu registro pode ser igualmente corrompido se empregados de empresas que fornecem informações sobre impostos ou falências a partir de registros judiciais ou do governo acidentalmente trocarem um dígito ou lerem erroneamente um documento.

As agências de crédito relatam ser impossível monitorar a precisão das 3,5 bilhões de informações sobre contas de crédito que recebem por mês. Elas precisam constantemente lidar com reclamações de consumidores que falsificam informações de credores ou utilizam empresas de reparo de crédito duvidosas que questionam todas as informações de relatório de crédito independentemente de sua validade. Para separar o joio do trigo, as agências de crédito norte-americanas utilizam um sistema automatizado chamado e-OSCAR (*Electronic Online Solution for Complete and Accurate Reporting*) para encaminhar aos credores discussões com os consumidores para verificação.

Se o seu relatório de crédito apresentou um erro, as agências não costumam contatar diretamente o cre-

dor para corrigir a informação. Como forma de economia, elas enviam as reclamações e evidências dos consumidores a um centro de processamento de dados gerenciado por uma empresa terceirizada. Essas terceirizadas rapidamente resumem cada reclamação em um pequeno comentário e associam a elas um código de dois dígitos a partir de uma lista de 26 opções. O código A3, por exemplo, representa 'pertence a outro indivíduo com nome semelhante'. Esses comentários costumam ser curtos demais para incluir o nome dos bancos necessários à compreensão da reclamação.

Embora o sistema corrija um grande número de erros (os dados são atualizados ou corrigidos em 72 por cento das reclamações), os consumidores dispõem de poucas opções em caso de falha. Aqueles que fazem uma segunda reclamação sem oferecer novas informações podem ter o pedido descartado como 'supérfluo'. Se o consumidor tentar, por conta própria, contatar o credor que cometeu o erro, os bancos não têm qualquer obrigação de investigar a reclamação — a menos que ela seja enviada por uma agência de crédito.

Fontes: Anne Kadet, "Why credit bureaus can't get it right". *Smart Money*, mar. 2009; "Credit report fix a headache". *Atlanta Journal-Constitution*, 14 jun. 2009; e Lucy Lazarony, "Your name can mess up your credit report". *Bankrate.com*, acesso em 01 jul. 2009.

PERGUNTAS SOBRE O ESTUDO DE CASO

1. Avalie o impacto sobre os negócios dos problemas de qualidade dos dados das agências de crédito. Considere agências de crédito, credores e indivíduos.
2. Existem questões éticas levantadas pelos problemas de qualidade dos dados das agências de crédito? Explique.
3. Analise os fatores humanos, organizacionais e tecnológicos responsáveis pelos problemas de qualidade dos dados das agências de crédito.
4. O que pode ser feito para resolver esses problemas?

Projetos práticos em SIG

Os projetos nesta seção oferecem experiência prática na análise de problemas de qualidade de dados, definição de padrões de dados que abarquem toda a empresa, criação de banco de dados para gestão de estoque e utilização da Web para pesquisa de bancos de dados on-line com informações sobre recursos empresariais internacionais.

Problemas de decisões gerenciais

1. A Emerson Process Management, fornecedora global de serviços e instrumentos de medida, analíticos e de monitoração com sede em Austin, Texas, projetou um novo data warehouse para análise das atividades de clientes e melhoria de serviços e marketing. Entretanto, o data warehouse estava repleto de dados imprecisos e redundantes. Os dados armazenados vinham de inúmeros sistemas de processamento de transações na Europa, na Ásia e em outras localidades pelo mundo. A equipe que projetou o data warehouse assumiu que as equipes de vendas de todas essas áreas informariam nomes e endereços de clientes da mesma maneira, independentemente de sua localização. Na verdade, as diferenças culturais, combinadas às complicações de absorção das empresas adquiridas pela Emerson, levaram a muitas formas de cotações, faturas, envios e outras informações. Avalie o possível impacto empresarial desses problemas de qualidade de dados. Quais decisões e medidas precisam ser tomadas para que se chegue a uma solução?

2. Sua empresa de abastecimento industrial deseja criar um data warehouse no qual a gerência obtenha visão completa das informações críticas de vendas de toda a empresa para identificar os produtos mais vendidos em áreas geográficas específicas, os principais clientes e as tendências de vendas. Suas informações sobre vendas e produtos estão armazenadas em diversos sistemas diferentes: um sistema de vendas por divisão rodado em um servidor UNIX, e um sistema de vendas corporativo rodado em um mainframe IBM. Você gostaria de criar um formato padrão único que consolide esses dados de ambos os sistemas. O seguinte formato foi proposto:

ID_PRODUTO	DESCRIÇÃO_PRODUTO	CUSTO_UNITÁRIO	UNIDADES_VENDIDAS	REGIÃO_VENDAS	DIVISÃO	ID_CLIENTE
------------	-------------------	----------------	-------------------	---------------	---------	------------

As imagens a seguir exemplificam os arquivos dos dois sistemas que forneceriam os dados para o data warehouse:

SISTEMA CORPORATIVO DE VENDAS

ID_PRODUTO	DESCRIÇÃO_PRODUTO	CUSTO_UNITÁRIO	UNIDADES_VENDIDAS	REGIÃO_VENDAS	DIVISÃO
60231	Rolamento, 4"	5,28	900.245	Nordeste	Peças
85773	Unidade de montagem	12,45	992.111	Melo-oeste	Peças

SISTEMA DE VENDAS POR DIVISÃO - PEÇAS MECÂNICAS

NUM_PRODUTO	DESCRIÇÃO_PRODUTO	CUSTO_UNITÁRIO	UNIDADES_VENDIDAS	REGIÃO_VENDAS	ID_CLIENTE
60231	Rolamento de aço de 4"	5,28	900.245	NE	Anderson
85773	Unidade de montagem	12,45	992.111	MO	Kelly Industries

- Quais problemas organizacionais são criados pelo fato de esses dados não estarem em um único padrão?
- Seria fácil criar um banco de dados com um único formato que pudesse armazenar os dados de ambos os sistemas? Identifique os problemas a serem resolvidos.
- Os problemas deveriam ser solucionados por especialistas em bancos de dados ou pelos gerentes gerais da empresa? Explique.
- Para as informações no data warehouse dessa empresa, quem deve ter autoridade para finalizar um formato único padrão?

Aperfeiçoando a tomada de decisão: pesquisando recursos empresariais internacionais em bancos de dados on-line

Habilidades de software: Bancos de dados on-line

Habilidades organizacionais: pesquisar serviços para operações no exterior

Os usuários da Internet têm acesso a milhares de bancos de dados com informações sobre serviços e produtos em lugares remotos. Este projeto ajudará você a desenvolver sua habilidade de pesquisar bancos de dados on-line.

Sua empresa situa-se em Greensboro, Carolina do Norte, e fabrica móveis de escritório de vários tipos. Recentemente, conquistaram vários clientes na Austrália, e um estudo encomendado indica que, se estivessem presentes neste país, poderiam aumentar consideravelmente as vendas. Além disso, o estudo indica que seria ainda melhor se fabricassem vários de seus produtos localmente (na Austrália). Primeiro, é necessário abrir um escritório em Melbourne para firmar presença e, depois, é preciso começar a importar dos Estados Unidos. Poderão então pensar em começar a produzir localmente.

Em breve, você viajará à região para planejar a abertura do escritório e aproveitará para visitar organizações que possam ajudá-lo na operação. Precisa contar com pessoas ou empresas que ofereçam vários serviços necessários à abertura do escritório, o que inclui advogados, contadores, especialistas em importação e exportação, equipamento e suporte de telecomunicações e até mesmo instrutores que possam ajudá-lo a preparar seus futuros funcionários.

Comece procurando os conselhos do Departamento de Comércio dos Estados Unidos sobre a prática de negócios na Austrália. Em seguida, experimente os seguintes bancos de dados on-line para localizar empresas que gostaria de contatar durante a viagem: Australian Business Register (abr.gov.au), Australia Trade Now (australiatradenow.com) e Nationwide Business Directory of Australia (www.nationwide.com.au). Se necessário, use ferramentas de busca como Yahoo! e Google.

- Enumere as empresas que gostaria de contatar para uma entrevista durante a viagem, a fim de determinar se elas podem ajudá-lo com essas e outras funções que considera vitais para o estabelecimento do escritório.
- Dê notas para os bancos de dados que usou levando em conta a precisão de nome, a completude, a facilidade de uso e a utilidade geral.
- O que este exercício lhe diz sobre o projeto de bancos de dados?

Resumo

1. Como um banco de dados relacional organiza os dados? Compare essa abordagem com o banco de dados orientado a objetos. O banco de dados relacional é, atualmente, o método mais usado para organizar e manter dados em sistemas de informação. Ele disponibiliza os dados em tabelas bidimensionais, com linhas e colunas chamadas relações. Cada tabela contém dados sobre uma entidade e seus atributos. Cada linha representa um registro, e cada coluna representa um atributo ou campo. Cada tabela também contém um campo-chave, que confere identificação exclusiva a cada registro para posterior recuperação ou manipulação. O diagrama entidade/relacionamento descreve graficamente a relação entre as entidades (tabelas) em um banco de dados relacional. Um banco de dados relacional bem projetado não terá relacionamentos muitos-para-muitos, e todos os atributos de uma entidade específica se aplicarão apenas àquela entidade. O processo de fragmentar agrupamentos complexos de dados e simplificá-los, a fim de minimizar a redundância e as relações muitos-para-muitos inadequadas, é denominada normalização. Um banco de dados orientado a objeto armazena dados e procedimentos como se fossem objetos, podendo manipular tanto elementos de multimídia quanto caracteres e números.

2. Quais os princípios de um sistema de gestão de banco de dados? Um sistema de gestão de banco de dados (DBMS) é um software que permite a centralização e a gestão de dados, de maneira que as empresas disponham de uma única fonte consistente para todas as necessidades de informações. Um único banco de dados atende a múltiplas aplicações. A mais importante característica do DBMS é sua capacidade de separar as visões lógica e física dos dados. Entre os principais recursos de um DBMS estão um recurso de definição, outro de dicionário de dados e uma linguagem de manipulação dessas informações.

3. Quais as principais ferramentas e tecnologias para extrair informações de bancos de dados e, assim, melhorar o desempenho da empresa e a tomada de decisão? As empresas dispõem de poderosas ferramentas para analisar e acessar as informações dos bancos de dados. Um data

warehouse consolida dados atuais e históricos de vários sistemas operacionais diferentes em um banco de dados central, projetado para proporcionar relatórios e análises. Os data warehouses suportam a análise multidimensional de dados, também conhecida como processamento analítico on-line (OLAP). O OLAP representa os relacionamentos entre os dados como uma estrutura multidimensional que pode ser visualizada como cubos de dados uns dentro dos outros, o que permite análises de dados mais sofisticadas. O data mining avalia grandes quantidades de dados, incluindo o conteúdo dos data warehouses, a fim de encontrar padrões e regras utilizados para prever o comportamento futuro e orientar a tomada de decisão. Ferramentas de mineração de texto ajudam as empresas a analisar grandes conjuntos de dados formados por textos não estruturados. Ferramentas de mineração na Web concentram-se na análise de parâmetros e informações úteis na World Wide Web a partir da avaliação da estrutura dos sites, das atividades dos usuários na Web e dos conteúdos das páginas. Bancos de dados convencionais podem ser conectados à Web ou a uma interface da Web para facilitar o acesso do usuário aos dados internos da organização.

4. Qual o papel da política de informação e da gestão de dados na gestão dos recursos de dados organizacionais? Desenvolver um ambiente de banco de dados exige políticas e procedimentos para gerenciar os dados organizacionais, assim como um bom modelo de dados e uma boa tecnologia de banco de dados. Uma política de informação formal regula a manutenção, a distribuição e o uso da informação na organização. Em grandes corporações, uma função formal de administração de dados é responsável pela política de informação, assim como pelo planejamento de dados, desenvolvimento do dicionário de dados e monitoria do uso dos dados da empresa.

5. Por que é importante para a empresa garantir a qualidade dos dados? Dados inexatos, incompletos ou inconsistentes criam sérios problemas financeiros e operacionais para as empresas, porque podem gerar imprecisões no preço de produtos, em contas de clientes e em informações de estoque, além de conduzir a decisões