

Qual o Papel de um DBA?

Saiba mais sobre o que faz um administrador de banco de dados e como se tornar um

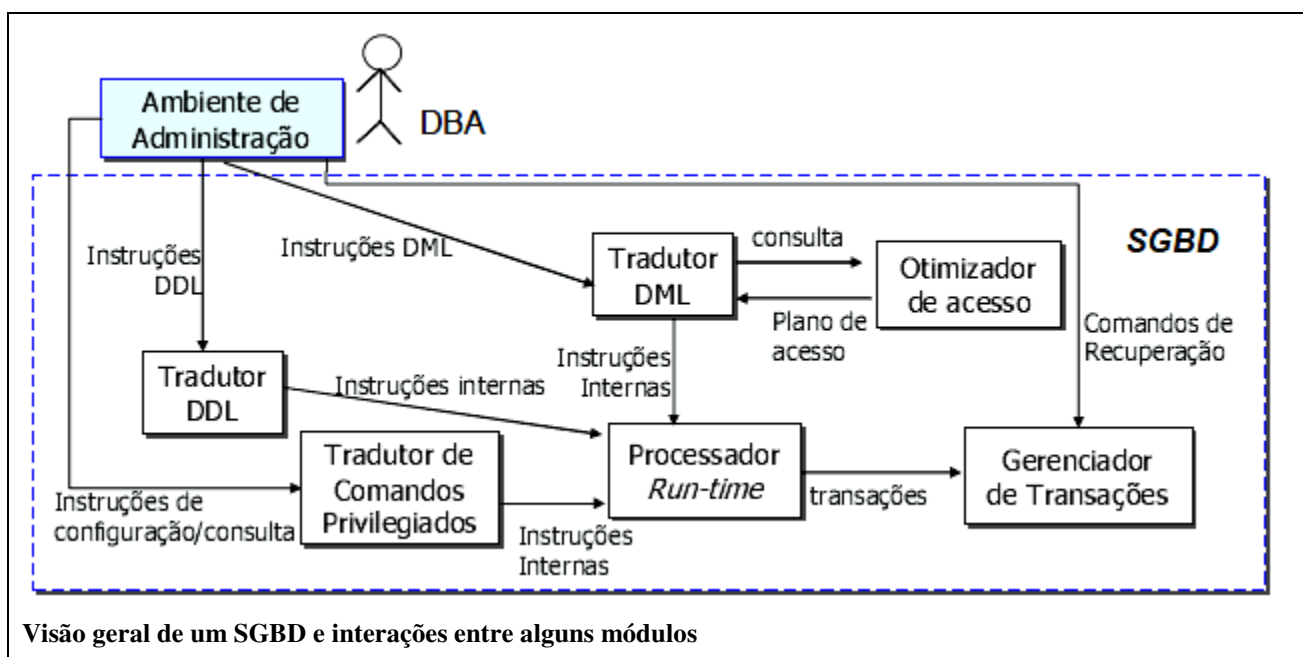
Carina Friedrich Dorneles, dorneles@upf.br, Universidade de Passo Fundo (UPF)

Ronaldo dos Santos Mello, ronaldo@inf.ufsc.br, Universidade Federal de Santa Catarina (UFSC)

O salário é atrativo, o assunto é fascinante, as ferramentas disponíveis são inúmeras, mas... a responsabilidade é altíssima! Neste artigo, procuramos responder à pergunta “Qual o papel de um DBA?” (Database Administrator, ou Administrador de Banco de Dados). Para isso, são apresentadas as atividades mais comuns exercidas por um DBA, suas responsabilidades e habilidades. Mostramos também as implicações de se tornar um profissional na área, apresentando alguns efeitos de uma má administração. Por fim, listamos algumas áreas interessantes de atuação e apontamos algumas dicas de como se tornar um DBA.

Quando aprendemos Banco de Dados na Graduação, já nas primeiras aulas o professor nos mostra conceitos, históricos, a importância dos dados para um sistema informatizado, a arquitetura de um Sistema de Gerência de Banco de Dados (SGBD) e os atores envolvidos em cada atividade que um SGBD exige para funcionar. Um dos atores mais importantes, nesse contexto, é o DBA que “detém a responsabilidade técnica central pelos dados”. Grande esta responsabilidade, não? Pois é, se entrarmos a fundo nas atividades deste ator, veremos que ele carrega nas costas a obrigação de responder pelas consequências dos próprios atos, e muitas vezes, dos outros.

O SGBD não trabalha sozinho. Ele exige um profissional qualificado que controle a máquina implementada dentro dele. Se olharmos a anatomia de um SGBD, veremos que ele possui inúmeros módulos que interagem entre si, como mostra a figura a seguir. No entanto, a configuração para que esta máquina funcione, da melhor maneira possível, com o máximo de eficiência e eficácia, é realizada pelo DBA. A figura a seguir apresenta os principais módulos de um SGBD, com destaque para algumas das atividades de responsabilidade do DBA.



Visão geral de um SGBD e interações entre alguns módulos

Atividades mais comuns exercidas

Como já comentado, o DBA tem um papel importante para o bom funcionamento de um sistema corporativo. Mas, quais são as atividades que, de fato, ele deve executar? Na prática, são inúmeras, desde o contato com o usuário final da aplicação até aquele “palpite” na hora de trocar o servidor do Banco de Dados. No entanto, podemos listar algumas tarefas que são de sua responsabilidade e que normalmente são exercidas *somente* por ele.

Instalação e configuração do servidor. A primeira tarefa do DBA é a instalação e configuração do servidor de banco de dados. Para que esta tarefa seja realizada com sucesso, é muito importante que o DBA tenha conhecimentos avançados do SGBD que está sendo instalado. Esta tarefa envolve atividades que vão desde o planejamento de *hardware* e sistema operacional, questões de segurança, licença de uso, até configurações básicas tais como controle de codificação de caracteres (definição do *encoding*), nome do banco de dados, número máximo de transações concorrentes permitidas, *timeout*, entre outros.

Definição do esquema. Esta atividade inicia a implementação de um banco de dados em um SGBD, correspondendo à definição da estrutura dos registros de dados do banco de dados. Neste ponto, o DBA recebe do analista de banco de dados (DA ou *Data Analyst* – profissional responsável por identificar os requisitos de dados da corporação) a modelagem conceitual (um diagrama entidade-relacionamento ou um diagrama de classes UML, por exemplo) e o converte para uma modelagem lógica (relacional, objeto-relacional, XML, ou outra) compatível com o modelo implementado pelo SGBD utilizado. Em alguns ambientes de negócio, o DBA acumula a função de DA e, portanto, deve ter conhecimento fundamental de análise de *software* e conhecimento aprofundado de modelagem de dados.

Definição da estrutura de armazenamento. Aqui o DBA inicia o processo de construção das estruturas internas do SGBD, tais como criação de índices e métodos de acesso. Existem também considerações a serem feitas a respeito de desempenho, controle de concorrência e requisitos de espaço de armazenamento (tamanho dos atributos de tabelas e tipos associados, dentre outros). Decisões como “armazeno figuras em atributos da tabela ou no sistema de arquivos do servidor de banco de dados?” são feitas durante a execução desta atividade. Muitas destas atividades necessitam de um profundo conhecimento da linguagem SQL para a definição de dados (SQL/DDDL: *Data Definition Language*), e dos tipos permitidos em cada SGBD.

Concessão de autorização de acesso aos dados. Você já deve ter acessado algum sistema Web em que se fazia necessária a autenticação de usuários (a página Web da rodoviária de Porto Alegre, um ambiente de Intranet de alguma empresa ou Universidade, por exemplo). Pois bem, estas permissões são normalmente concedidas ou revogadas pelo DBA. Como ele é responsável pela criação do esquema dos dados e pela definição da estrutura de armazenamento, ele também deve se responsabilizar por “quem pode acessar o quê”. Dependendo do sistema de aplicação e das necessidades de cada empresa, alguns usuários só possuem privilégio para consultar certos dados, enquanto outros podem também atualizar dados. Um bom conhecimento da SQL/DDDL também se faz necessário.

Especificação das restrições de integridade. Não basta apenas definir o esquema dos dados de um banco de dados. O DBA deve garantir que o SGBD mantenha valores válidos para estes dados e transições de valor igualmente válidas. Questões como “um determinado atributo de uma tabela pode ser nulo?”, “o salário de um funcionário pode ser reduzido?” ou ainda “é admissível que duas pessoas sejam associadas ao mesmo e-mail? Ou ele deve ser único por pessoa?” devem ser respondidas pela corporação e configuradas no banco de dados pelo DBA. Além disso, esta atividade envolve definições importantes de restrições de integridade referencial: “uma coluna definida como chave estrangeira não pode referenciar um valor de chave primária que ainda não existe”. Novamente: conhecimento da SQL/DDDL e da SQL/DML (*Data Manipulation Language*) é fundamental!

Monitoramento de desempenho. A tarefa de acompanhar o desempenho do SGBD é uma das mais complexas de um DBA, pois, além de efetuar o monitoramento, é necessário que ele esteja preparado para tomar decisões importantes sobre “o que fazer se o desempenho não estiver bom”. Muitas vezes, isso implica diretamente na situação financeira de uma corporação, pois o DBA deve escolher entre “adquirir novo *hardware*” ou simplesmente “ajustar o desempenho do SGBD” (veja a próxima atividade).

Melhoria de desempenho. A melhoria de desempenho (*performance*) envolve diversas sub-tarefas, desde ajustes específicos de cada SGBD, tais como escolha do melhor plano de execução de uma consulta, definição de visões (*views*) para a geração de relatórios complexos, seleção de um mecanismo de bufferização de dados adequado e redefinição de índices, até questões relacionadas à modelagem e projeto do banco de dados. Uma modelagem, ou projeto, mal definida gera uma estrutura interna ineficiente em termos de acesso, onde vários problemas de desempenho podem ser detectados. A utilização adequada da infra-estrutura (memória, unidades de armazenamento, processadores, entre outros) é outra questão importantíssima quando o assunto é desempenho.

Definição de estratégias de backup. A atividade de *backup* é uma tarefa crucial em todos os ambientes corporativos, pois qualquer empresa pode ser vítima de algum acidente, ou incidente, que pode causar perda de dados. Uma vez que DBAs são responsáveis pelos dados destas corporações, é natural que seja deles a decisão pela melhor estratégia de *backup* em cada caso. De forma geral, pode-se definir duas estratégias: uma cópia completa do banco de dados, ou uma cópia consistente para leitura e sem apresentar as transações ativas. Vale salientar que muitos SGBDs oferecem ao DBA ferramentas para controle eficiente de procedimentos de *backup*, como por exemplo, *backups* incrementais e *backups* on-line, que evitam a degradação de desempenho do banco de dados quando estão executando.

Definição de estratégias de recuperação de dados. A tarefa de recuperação de dados requer um SGBD tolerante a falhas, ou seja, um SGBD que seja capaz de conduzir o banco de dados a um estado passado consistente, após a ocorrência de uma falha que o deixou em um estado inconsistente. Esta “tolerância a falhas” muitas vezes se baseia em redundância de dados. Um tipo de técnica de recuperação muito usada é a baseada em *logs*, que são arquivos que registram seqüencialmente as atualizações no banco de dados. *Logs* normalmente são mantidos em cópias em memória secundária (disco ou fita, entre outros). O DBA deve conhecer detalhadamente cada técnica implementada nos SGBDs a fim de definir a melhor alternativa no caso de uma falha.

Modificação de esquema. O esquema de um banco de dados pode evoluir com o passar do tempo, seja devido a modificações nos requisitos de dados da corporação, seja para melhoria de desempenho. Em alguns casos, são necessárias noções de engenharia reversa de banco de dados, onde se obtém a modelagem conceitual do banco de dados a partir da sua implementação em um SGBD, visando a redefinição de requisitos ou de estratégias de mapeamento para esquemas lógicos mais eficientes. Vale ressaltar que a modificação do esquema de um banco de dados deve ser realizada com cautela, de modo a não tornar inválidos dados eventualmente já cadastrados, como por exemplo, a redefinição de uma chave primária ou a alteração do tipo de dado ou de uma restrição de integridade de um atributo em uma tabela.

Reorganização física. A tarefa de organização física do banco de dados trata da disposição e localização física dos dados nos dispositivos de armazenamento. Como um banco de dados é composto por vários arquivos relacionados entre si, é necessário que eles estejam armazenados em um local físico. A localização desses arquivos influencia a forma como são acessados, como são inseridos ou alterados, o tempo de resposta de cada consulta, a segurança e o custo de processamento. Os arquivos do banco de dados controlados por um SGBD podem ser centralizados ou distribuídos. O DBA deve organizar os dados fisicamente de forma a usar de maneira eficiente os recursos de armazenamento e processamento. Por exemplo, dados distribuídos em vários computadores controlados pelo SGBD

garantem uma maior disponibilidade. Entretanto, se estes dados estiverem replicados, haverá um maior custo para mantê-los consistentes quando da ocorrência de atualizações. Em alguns casos, quando o desempenho do banco não está satisfatória, é necessária a execução de uma reorganização física dos dados.

Conseqüências de uma administração ruim

Para termos a noção das responsabilidades de um DBA, vamos analisar algumas conseqüências de uma má administração de um banco de dados. Como já mencionado, uma má administração pode começar na fase de projeto do banco de dados. Uma modelagem conceitual de dados mal construída pode gerar conseqüências desastrosas no banco de dados. Isso inclui a identificação, ou não, de uma entidade, ou mesmo a definição errônea de uma cardinalidade. Podemos ter, por exemplo, o caso de uma corporação em que é necessário armazenar clientes e suas profissões. Se o banco de dados for projetado de forma a possibilitar que um cliente possa ter várias profissões e que uma profissão possa estar associada a vários clientes, teremos uma tabela associativa no banco de dados. No entanto, a criação de uma tabela associativa implica a criação de mais chaves primárias, mais índices, e consultas com acesso a um maior número de tabelas, o que pode causar, dependendo do volume de dados, uma baixa considerável no desempenho. Portanto, todo projeto deve ser revisto, sempre! A teoria da Normalização e suas formas normais são grandes aliadas para esta atividade.

Perda de dados é outra conseqüência de uma má administração, que pode acontecer em situações diversas, tais como uma falha que interrompe a execução do SGBD ou a falta de definição de *backups*. A execução programada de *backups*, em uma frequência adequada, é uma tarefa que evita muitos problemas futuros. Além disso, o DBA deve estar preparado, também, para recuperar o banco de dados na ocorrência de uma falha. Para isso, é muito importante que ele tenha em mãos alguma ferramenta a qual o DBA conheça muito bem, a fim de não se deparar com problemas no uso da ferramenta em um momento em que a falha ocorrer.

Finalmente, o baixo desempenho do banco de dados é outra conseqüência que, em certos casos, pode ser decorrente de uma má administração. No entanto, é importante estar atento: nem sempre o baixo desempenho é resultado de algum problema no banco de dados. Outros fatores, como conexões ou configurações de rede, bem como configuração de *hardware*, podem estar associados ao problema. Para diagnosticar o banco de dados, é muito útil a utilização de ferramentas de monitoramento que, se usadas de forma preventiva, muitas vezes evitam que o desempenho caia a níveis muito baixos. O conhecimento detalhado de tais ferramentas também é muito importante, pois uma degradação acentuada de desempenho em função de ausência de monitoramento pelo DBA pode inviabilizar a execução de operações vitais da corporação.

Como se tornar um DBA e onde atuar

Um bom DBA tem um caminho longo no currículo. Além do conhecimento da área de Banco de Dados, adquirido em bons cursos de graduação (e eventualmente de pós-graduação), um bom diferencial (ou um bom ponta-pé inicial) pode ser uma certificação (veja a lista de Recursos). Grandes corporações geralmente trabalham com SGBDs robustos e você pode adquirir uma certificação destes SGBDs.

Treinamentos realizados com certa frequência também são relevantes, principalmente quando um determinado SGBD lança um novo *release* que provê suporte, por exemplo, a um novo formato de dados do qual o DBA não tem conhecimento. Em alguns casos, novos SGBDs que lidam com a gerência específica de um formato de dados, como é o caso do formato XML nos últimos anos, tornam-se disponíveis comercialmente. É importante que o DBA esteja atento a estes fatos! Treinamentos podem ser feitos tanto em cursos de especialização quanto em empresas do ramo.

A área de atuação de um DBA é ampla e engloba desde atividades em empresas públicas ou privadas de médio e grande porte até consultorias em projetos financiados pelo governo, pólos de *software*, parques tecnológicos e grandes Universidades.

Concluindo

Apesar de todas as atividades específicas mencionadas, é importante observar que o DBA não deve se deter apenas aos conhecimentos específicos de Banco de Dados. Como os SGBDs são sistemas complexos, configuráveis de acordo com as necessidades e infra-estrutura disponibilizadas, o profissional de Banco de Dados deve ter conhecimentos relativamente bons sobre sistemas operacionais, redes, formas de armazenamento, desenvolvimento de *software* e, nos dias de hoje, até mesmo alguma noção sobre alto desempenho. Tudo isso é necessário para a perfeita harmonia entre o DBA e toda a equipe de um departamento de TI.

Lembre-se que o DBA é o maior responsável pelos dados armazenados em um banco de dados. Estes dados são vitais para o correto funcionamento dos sistemas computacionais de uma corporação! Assim, quando alguma coisa está errada no servidor de banco de dados, o DBA precisa saber qual a melhor maneira de resolver o problema da forma mais rápida possível, realizando correções sem agravar a situação ou perder dados.

Se você se interessou pelo assunto e quer mais detalhes sobre o tema, é interessante consultar livros especializados. Como a execução das tarefas mencionadas depende de ferramentas específicas, é normal que os livros sejam dirigidos a um SGBD em particular. No entanto, você pode matar sua curiosidade mais geral com o livro mencionado na seção Recursos.

Recursos

Livro de referência: Craig S. Mullins, *Database Administration: The Complete Guide to Practices and Procedures*. Addison-Wesley Professional, 2002.

Você pode encontrar mais informações sobre certificações nos sites específicos de SGBDs, tais como:

- <http://education.oracle.com>
- <http://www.ibmdatabasemag.com/DB2>
- <http://www.microsoft.com/learning>
- <http://www.postgresqlcertification.org>
- <http://www.mysql.com/certification>

Sobre os autores



Carina Friedrich Dorneles é doutora em Ciência da Computação e professora da Universidade de Passo Fundo (UPF), nos cursos de Ciência da Computação e Sistemas de Informação. Sua área de ensino e pesquisa é Banco de Dados. Foi orientadora de trabalhos de iniciação científica, graduação e pós-graduação nas diversas sub-áreas de banco de dados, tais como planos de consultas, decomposição de consultas, entre outros. Já ministrou cursos e palestras na área, além da publicação de artigos científicos. Seus tópicos atuais de pesquisa envolvem: integração de dados, gerenciamento e manipulação de dados XML, casamento de aproximado de dados (com funções de similaridade).



Ronaldo dos Santos Mello é doutor em Ciência da Computação e professor adjunto da Universidade Federal de Santa Catarina (UFSC). Atua em Ciência da Computação e sua área de ensino e pesquisa é Banco de Dados. Foi orientador de diversos trabalhos de iniciação científica, graduação e pós-graduação em nível de especialização e mestrado. Coordena o Grupo de pesquisa em Banco de Dados da UFSC (GBD-UFSC). Já ministrou vários cursos e palestras na área de Banco de Dados, além da publicação de artigos científicos. Seus tópicos atuais de pesquisa envolvem: gerência de dados XML, integração de dados e bancos de dados não-convencionais.