



## ADMINISTRANDO DADOS e BANCOS de DADOS

---

Aqui vemos, o papel de administrador de dados principalmente em comparação com a tarefa de administração de banco de dados (DBA versus DA). Em seguida, apresentaremos o trabalho atribuído aos administradores de dados, citando metodologias e cuidados normalmente adotados.

### DBA VERSUS DA

O administrador de dados não é um dos vários usuários de um sistema de banco de dados como indicado na literatura normalmente referenciada. Em geral, existem desde usuários que são responsáveis pela gerência, acesso e desempenho do banco (os DBA - *DataBase Administrators*) até aqueles usuários que submetem requisições ao banco eventualmente, passando pelos desenvolvedores de programas de aplicação e usuários especializados.

No caso de empresas que possuem (ou podem vir a ter) diversos sistemas de bancos de dados - como o IBGE, é preciso ter vários administradores de bancos de dados (não necessariamente um para cada banco). Porém, é natural nessas situações que as definições e o processamento de dados originalmente idênticos ocorram de forma não uniforme, prejudicando uma visão global única desses dados, como seria o esperado. Mais ainda, quase não ocorre o reaproveitamento de trabalho realizado anteriormente na execução de novas tarefas, principalmente no momento da modelagem conceitual, onde o conhecimento da aplicação é de fundamental importância.

Assim surge a função de administrador de dados (DA - *Data Administrator*) na empresa, que herda algumas das responsabilidades comumente atribuídas a um DBA. Espera-se que esse novo usuário tenha não somente a visão dos conceitos e projeto de um banco de dados específico, como também uma compreensão e conhecimento dos diversos outros sistemas de banco de dados da empresa.

As funções normalmente atribuídas a um DA são:

- Levantar os requisitos funcionais para o banco de dados;
- Modelar conceitualmente o banco de dados;
- Especificar as regras de negócio das aplicações;
- Definir padrões de nomes para conceitos e variáveis, e
- Determinar normas de incorporação e manuseio dos dados.

---

 **Já o DBA tem como responsabilidades:**

- Mapear o modelo conceitual no modelo lógico;
- Realizar o projeto físico do banco de dados;
- Criar usuários, definir visões e permissões, além de regras de integridade ;
- Controlar os processos de *back-up* e recuperação, e
- Garantir o bom desempenho no acesso e manuseio do banco pelos usuários.

Ou seja, enquanto o DBA é o responsável pelo funcionamento adequado do sistema de banco de dados implementado na empresa, independentemente da semântica atribuída aos dados, o DA é o conhecedor das informações representadas no banco e deve ter a preocupação com a integração das informações entre bancos de dados distintos sempre que possível, fornecendo ao DBA a visão dos dados, sem preocupações com armazenamento e acesso posteriores.

O problema é que, na prática, essa separação de tarefas e funções nem sempre ocorre de forma transparente, dificultando sua implementação. Muitos DA assumem tarefas de DBA e vice-versa. Além disso, mesmo quando há uma separação de funções claramente especificada, podem ocorrer atritos sobre o que cada um, DA e DBA, esperam de um sistema. Enquanto o DA normalmente tem maior conhecimento da aplicação e da importância do sistema e dos dados para a empresa, o DBA tem por preocupação garantir um bom desempenho e a integridade do banco de dados a todo custo.

Os dois objetivos nem sempre são compatíveis relativamente ao grau exigido por cada um e normalmente o *tuning* ótimo para o DBA é atingido relaxando-se definições especificadas pelo DA ou vice-versa. Apenas para citar um exemplo simples, o controle de restrições de integridade semânticas, em geral orientadas à cada aplicação, pode implicar em procedimentos internos que reduzam o desempenho do sistema quando ocorrem atualizações. Enquanto para o DA esse controle é fundamental e deveria ser garantido pelo SGBD a todo momento, o DBA poderia imaginar um controle *off-line* com verificação das condições, e eventual correção, posteriormente à execução das transações.

Além disso, é também comum que a implementação do banco de dados pelo DBA seja retardada por falta de definições especificadas pelo DA ou mesmo que o DBA leve um tempo maior do que o esperado pelo DA em tornar operacional o banco de dados. Esses pequenos problemas são esperados visto que a tarefa do usuário principal do banco de dados é dividida em duas partes e executada por dois usuários distintos, sem estar definido qual deles teria a prioridade em caso de discórdia quanto ao detalhamento do sistema.

Normalmente o DA tem maior poder político, pelo próprio fato de ter conhecimento dos detalhes da aplicação e ser, conseqüentemente, aquele que superiores hierárquicos na empresa procuram para definir novas aplicações. Porém, o DBA é aquele que implementa as idéias concebidas pelo DA e tem o domínio do sistema, inclusive definindo quem pode acessar o banco de dados e com que direitos.

Interessante constatar que muitas vezes diz-se que os DAs fazem o trabalho "que ninguém quer fazer" e que os DBAs não quiseram assumir. Porém, muitas vezes acontece exatamente o contrário e existem aqueles usuários desenvolvedores do sistema que não querem se aproximar de aspectos puramente técnicos e que se interessam exatamente no processo de abstração e modelagem conceitual.

Também cabe constatar que o trabalho de DA não necessariamente precisa ser executado por algum especialista em informática, já que o fundamental é que seja alguém que tenha o conhecimento da aplicação em si. Entretanto, muitas das metodologias utilizadas por um administrador de dados são oriundas da área de informática, como o projeto conceitual do banco de dados. Logo, é comum que os DAs tenham formação básica em ciência da computação pois caso contrário podem não ser capazes de se responsabilizarem sozinhos por suas tarefas.

As experiências bem sucedidas de sistemas de informação mantidos simultaneamente por DAs e DBAs são quase todas devido ora a ambos estarem integrados em um mesmo departamento da empresa e, conseqüentemente, terem uma chefia comum que pode arbitrar as eventuais divergências, ora ao fato de cada um ser especializado em suas tarefas e existirem processos operacionais internos à empresa, sistematizados, que determinem a forma como devem interagir durante o processo de desenvolvimento das aplicações.

Independente de cada situação particular, não há dúvida de que a definição, alocação e divulgação das tarefas ao DA e ao DBA anteriormente ao desenvolvimento dos sistemas de banco de dados da empresa é de fundamental importância para o sucesso de sua implementação e entrada em operação.

## **TAREFAS DE ADMINISTRAÇÃO DE DADOS**

As tarefas normalmente atribuídas aos administradores de dados podem ser classificadas em função do momento em que são aplicadas: durante o processo de concepção do banco de dados (tarefas estáticas) ou já quando o banco de dados está operacional (tarefas dinâmicas).

Como visto anteriormente, o processo de análise e definição do sistema é realizado pelo DA, que determina o escopo e aplicabilidade do sistema e executa o processo de abstração do mundo real para o sistema de informação a ser construído. Nesse processo é feita a modelagem conceitual do banco de dados a partir de um levantamento das necessidades e especificações da aplicação, modelagem esta complementada pelas regras de manutenção de integridade específicas ao negócio.

Além disso, o DA tem o papel fundamental de definir padrões de nomes para as variáveis definidas para o problema, de acordo com a política implementada na empresa e devido a existência de outros sistemas. Isto implica na necessidade do conhecimento das diversas aplicações corporativas, com a previsão de comparações de ítems semanticamente idênticos que não devem estar codificados diferentemente. Já com o banco de dados em operação, o DA deve determinar como deve ser realizada a incorporação de dados ao banco e também a manutenção dos dados e respectivos metadados.

Em um ambiente distribuído, é natural que co-existam vários DAs e DBAs em consequência dos diferentes sistemas existentes. Isto dificulta o processo de padronização de conceitos e nomes, visto que os dados estão particionados e os modelos utilizados seguem particularidades relativas a cada caso .

Assim, surge a proposta de um grupo de administração de dados para a discussão e normalização dos dados para a empresa como um todo, com metodologias e técnicas elaboradas e adaptadas às necessidades específicas da aplicação, dúvidas e divergências solucionadas e novos procedimentos encaminhados em comum acordo e atendendo as especificidades da empresa como um todo.

Entretanto, é importante que a administração de dados seja prioritariamente realizada localmente a cada banco de dados, deixando para o comitê de administração de dados apenas questões pendentes que, na pior das hipóteses, não têm um tratamento uniforme mas que podem ser resolvidas por consenso. Os sistemas de informação mais recentes contemplam ferramentas de auxílio tanto para o DBA como para o DA. Nesse último caso é importante que o sistema contenha o maior conjunto de controles possível quanto às restrições de integridade, principalmente quanto ao processo de incorporação de dados por parte dos diversos usuários autorizados para tal.

Enquanto que para as tarefas dinâmicas é perfeitamente viável que se realize o controle da qualidade das informações sob forma de um programa de computador, para as

tarefas estáticas o que existe normalmente é um conjunto de técnicas e metodologias bem definido, utilizado de maneira rigorosa e respeitando-se certos padrões convencionados na empresa. Porém, é possível que a modelagem conceitual seja ajudada por ferramentas CASE, que facilitem inclusive o posterior mapeamento do modelo conceitual para o modelo lógico. Aqui cabe observar que, em muitas experiências práticas de administração de dados, o referido mapeamento também é feito pelo DA, em conjunto com o DBA.

Uma das tarefas mais complexas de um DA é a de compatibilizar conceitos de áreas em princípio distintas. A interpretação dada a um conceito pode determinar variáveis importantes para alguns bancos de dados e para outros não. Isso pode acarretar problemas sérios de comparabilidade entre dados, sobre vários conjuntos de informações diversos ou mesmo entre as duas ou mais instâncias de um mesmo conjunto de atributos definido. A padronização de nomes não é então um mero formalismo com regras de derivação das variáveis mas sim um processo de interpretação de conceitos e definição de variáveis que leva em consideração aspectos relevantes para a empresa como um todo. Esse padrão vem restringir o processo de incorporação de dados ao banco, visando a manutenção das normas especificadas para todos os sistemas.

**DI/COBAD**