

SISTEMAS DE GERÊNCIA DE BANCO DE DADOS

Para entender os Sistemas Gerenciadores de Banco de Dados é importante conhecer alguns conceitos básicos. A primeira definição é relativa aos conceitos de dados e informação. **Dados** são fatos em sua forma primária, os quais podem ser armazenados, como, por exemplo: nome, telefone e endereço. Estes dados ou fatos organizados de maneira significativa e relacionados formam uma **informação**, como por exemplo: os dados das peças em estoque. Assim, é possível obter uma lista de peças em falta.

Sabendo o que são dados e informações é possível definir os **Bancos de Dados** como um conjunto de dados devidamente relacionados capazes de apresentar uma informação.

Os bancos de dados são utilizados em muitas aplicações, abrangendo praticamente todo o campo de programas de computador. Os bancos de dados são mecanismos de armazenamento preferencial para aplicações multiusuário, nas quais é necessário haver coordenação entre vários usuários. Dessa forma, outra definição para banco de dados é que é um conjunto de informações com uma estrutura regular, normalmente, mas não necessariamente, armazenado em algum formato de máquina legível para um computador.

Há uma grande variedade de bancos de dados, desde simples tabelas armazenadas em um único arquivo até um conjunto com milhões de registros armazenados em salas cheias de discos rígidos gerenciados por uma determinada aplicação.

Os dados armazenados em bancos de dados geralmente são apresentados em forma semelhante à uma planilha eletrônica, porém existem sistemas de gestão responsáveis por gerir o armazenamento, classificação, gestão da integridade e recuperação dos dados.

2.1 CARACTERÍSTICAS DE UM BANCO DE DADOS

Como vimos, um banco de dados é uma coleção lógica coerente de dados com um significado inerente. Assim, uma disposição desordenada dos dados não pode ser referenciada como um banco de dados;

Um banco de dados deve ser projetado, construído e populado com dados para um propósito específico;

Deve possuir um conjunto pré-definido de usuários e aplicações;

Representar algum aspecto do mundo real, o qual é chamado de “mini-mundo”. Qualquer alteração efetuada no mini-mundo é automaticamente refletida no banco de dados.

2.2 MODELOS DE BANCO DE DADOS

2.2.1 Modelo Hierárquico ou de árvore

É aquele no qual os dados estão organizados de cima para baixo ou estrutura de árvore invertida. Por exemplo, os dados sobre um projeto para uma empresa podem seguir este tipo de modelo. Este método de ligação é semelhante à relação entre pais e filhos: a criança não existirá sem os pais. É o que mais bem se adapta a situações nas quais as relações lógicas entre os dados podem ser representadas com a abordagem (um-para-muitos).

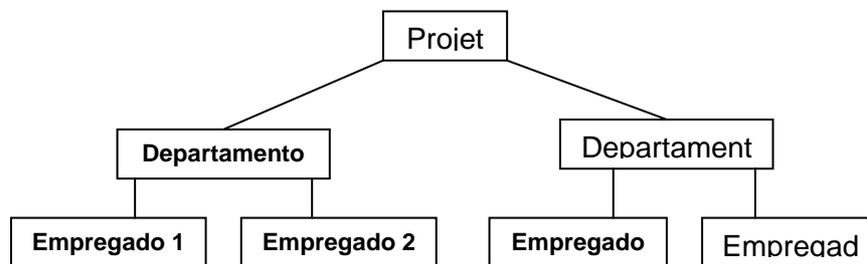


Figura 3 - Modelo Hierárquico

2.2.2 Modelo em Rede

Um modelo em rede é uma extensão do modelo hierárquico. Em vez de se terem apenas vários níveis de relações um-para-muitos, o modelo em rede é uma relação membro-proprietário, na qual um membro pode ter muitos proprietários. Nesse modelo, há frequentemente mais de um caminho pelo qual um determinado elemento de dado pode ser acessado.

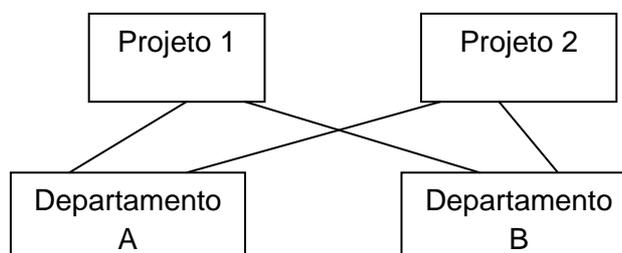


Figura 4 - Modelo em Rede

2.2.3 Modelo Relacional

Os modelos relacionais se tornaram os mais populares. A finalidade global deste modelo é descrever o dado usando um formato tabular padrão (todos os elementos são localizados em tabelas bidimensionais). As tabelas organizam os dados em linhas e colunas, simplificando o acesso e a manipulação dos dados.

Uma vez colocados os dados no Banco de Dados Relacional, pode-se fazer perguntas e manipular dados utilizando as operações da álgebra relacional. As manipulações básicas de dados incluem a sua seleção, projeção e agrupamento. Outros aspectos deste modelo serão abordados nos capítulos seguintes.

2.3 SISTEMA DE GERENCIADOR DE BANCO DE DADOS (SGBD)

Um banco de dados pode ser criado e mantido por um conjunto de aplicações desenvolvidas especialmente para esta tarefa ou por um Sistema Gerenciador de Banco de Dados. **Este é um software com recursos específicos para facilitar a manipulação das informações dos bancos de dados e o desenvolvimento de programas aplicativos.**

Outro conceito interessante sobre SGBD pode ser encontrado no Wikipédia:

...o conjunto de programas de computador (softwares) responsáveis pelo gerenciamento de uma base de dados.

Os Sistemas Gerenciadores de Bancos de Dados surgiram para atender à necessidade de armazenamento e de recuperação de grandes volumes de informações, propiciando um ambiente seguro e adequado.

Seu **principal objetivo** é retirar da aplicação cliente a responsabilidade de gerenciar o acesso, manipulação e organização dos dados. Estes sistemas disponibilizam uma interface para que os seus clientes possam incluir, alterar ou consultar dados.

2.3.1 Funções Básicas de um Sistema Gerenciador de Banco de Dados

Como todo e qualquer aplicativo os sistemas gerenciadores de bancos de dados dispõem de funções que são compartilhadas por qualquer tipo de sistema gerenciador de banco de dados, independente de seu fabricante. Dentre elas, destacam-se:

- Proporcionar maior abstração, isolando o usuário dos pormenores internos de como os dados estão armazenados;
- Rapidez no acesso às informações;
- Proporcionar independência dos dados em relação aos programas de recuperação, cuja estrutura física de armazenamento independe da estratégia de acesso;
- Compartilhar a base de dados entre vários aplicativos, em que diferentes tipos de interfaces atendem às necessidades de distintos usuários;

- Maior facilidade de realizar cópias de segurança;
- Proporcionar a comunicação diretamente com um software ou servidor;
- Proporcionar integridade dos dados.

2.3.2 Arquitetura de um Sistema Gerenciador de Banco de Dados

Atualmente, existem várias tendências para arquitetura de Banco de Dados. A arquitetura mais conhecida é a ANSI/SPARC. Esta arquitetura apresenta-se fundamentada em três níveis onde cada um desses níveis corresponde às abstrações dos dados armazenados no banco de dados.

Os esquemas são definidos como:

a) Nível externo: ou esquema de visão, o qual descreve o modo pelo qual os dados são vistos pelos usuários do sistema gerenciador de banco de dados. Cada visão descreve quais porções do banco de dados um usuário ou grupo de usuários terá acesso.

b) Nível conceitual: ou esquema conceitual, o qual descreve a estrutura do banco de dados como um todo, inclusive como os dados se relacionam. É uma descrição global do banco de dados que não fornece detalhes do modo como os dados estão fisicamente armazenados.

c) Nível interno ou físico: ou esquema interno, o qual descreve a estrutura de armazenamento físico do banco de dados. Utiliza um modelo de dados e descreve detalhadamente os dados armazenados e os caminhos de acesso ao banco de dados. É a descrição mais próxima de como os dados serão armazenados.

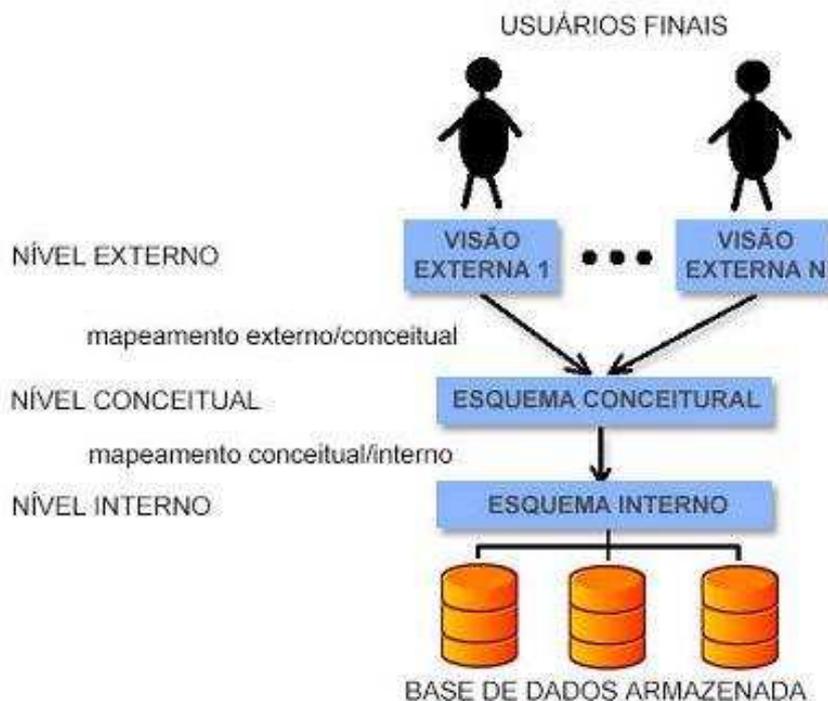


Figura 5 - Arquitetura de SGBD

Partindo da arquitetura proposta é importante conhecer alguns conceitos diretamente relacionados:

a) Abstração de Dados: Um Sistema Gerenciador de Banco de Dados é composto de uma coleção de arquivos inter-relacionados e de um conjunto de programas que permitem aos usuários realizarem acesso aos arquivos e modificá-los. O grande objetivo de um sistema de banco de dados é prover aos usuários uma **visão abstrata dos dados**. Isto é, **o sistema omite certos detalhes de como os dados são armazenados e mantidos**. No entanto, para que o sistema possa ser utilizado, os dados devem ser buscados de forma eficiente. Este conceito tem direcionado o projeto de estrutura de dados complexas para a representação de dados em um banco de dados. Uma vez que muitos dos usuários de banco de dados não são treinados para computação, a complexidade está escondida desses usuários através de **diversos níveis de abstração** que simplificam a interação do usuário com o sistema.

b) Nível de Visões: O mais alto nível de abstração descreve apenas parte do banco de dados. Apesar do uso de estruturas mais simples do que no nível conceitual, alguma complexidade perdura devido ao grande tamanho do banco de dados. Muitos usuários do sistema de banco de dados não estarão interessados em todas as informações. Em vez disso precisam de apenas uma parte do banco de dados. O nível de abstração das visões de dados é definido para simplificar esta interação com o sistema, que pode fornecer muitas visões para o mesmo banco de dados.

Dentro deste conceito, no **nível físico** um registro *cliente*, *conta* ou *funcionário* pode ser descrito como um bloco de posições de armazenamento consecutivo (por exemplo, palavras ou bytes). No **nível conceitual**, cada registro destes é descrito por uma definição de tipo e pelo inter-relacionamento entre esses tipos de registros. No **nível de visões**, diversas visões do banco de dados são definidas, por exemplo: os contadores de um banco vêem apenas a parte do banco de dados que possui informações sobre contas dos clientes. Eles não podem ter acesso a informações que se referem a salário dos funcionários.

c) Independência de Dados: Conhecemos três níveis de abstração pelos quais o banco de dados pode ser visto. A habilidade de modificar a definição de um esquema em um nível sem afetar a definição de esquema num nível mais alto é chamada de *independência de dados*.

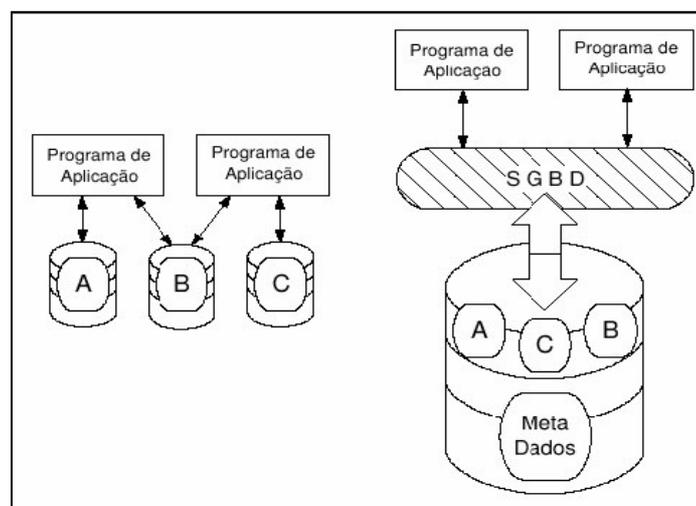


Figura 6 – Independência de dados

Existem dois níveis de independência dos dados:

- **Independência Física de Dados:** É a habilidade de modificar o esquema físico sem a necessidade de reescrever os programas aplicativos. As modificações no nível físico são ocasionalmente necessárias para melhorar o desempenho;

- **Independência Lógica de Dados:** É a habilidade de modificar o esquema conceitual sem a necessidade de reescrever os programas aplicativos. As modificações no nível conceitual são necessárias quando a estrutura lógica do banco de dados é alterada (por exemplo, a adição de contas de bolsas de mercado num sistema bancário).

A independência lógica dos dados é mais difícil de ser alcançada do que a independência física, porém os programas são bastante dependentes da estrutura lógica dos dados que eles acessam.

Este conceito de independência dos dados é similar em muitos aspectos ao conceito de tipos abstratos de dados em modernas linguagens de programação. Ambos escondem detalhes de implementação do usuário. Isto permite ao usuário concentrar-se na estrutura geral em vez de detalhes de baixo nível de implementação.

d) Dicionário de Dados: Um dicionário de dados é uma **coleção de metadados** que contém definições e representações de elementos de dados. Dentro do contexto de sistemas gerenciadores de bancos de dados, um dicionário de dados é um grupo de tabelas, habilitadas apenas para leitura ou consulta, ou seja, é uma base de dados, propriamente dita, que entre outras coisas, mantém as seguintes informações:

- Definição precisa sobre elementos de dados
- Perfis de usuários, papéis e privilégios
- Descrição de objetos
- Integridade de restrições
- *Stored procedures*¹ e gatilhos
- Estrutura geral da base de dados
- Informação de verificação
- Alocações de espaço

Um dos benefícios de um dicionário de dados bem preparado é a consistência entre itens de dados através de diferentes tabelas. Por exemplo, diversas tabelas podem conter números de telefones. Utilizando uma definição de um dicionário de dados bem feito, o formato do campo 'número de telefone' definido com "(99)9999-9999" deverá ser obedecido em todas as tabelas que utilizarem esta informação.

¹ pequeno trecho de programa de computador, armazenado em um SGBD, que pode ser chamado freqüentemente por um programa principal.

Os dicionários de dados são gerados, normalmente, separados do Modelo de Dados visto que estes últimos costumam incluir complexos relacionamentos entre elementos de dados.

e) Modelo de Dados: Uma das principais características da abordagem de banco de dados é que fornece alguns níveis de abstração de dados omitindo ao usuário final detalhes de como estes dados são armazenados. Um modelo de dados é um conjunto de conceitos que podem ser utilizados para descrever a estrutura lógica e física de um banco de dados. Por estrutura podemos compreender o tipo dos dados, os relacionamentos e as restrições que podem recair sobre os dados.

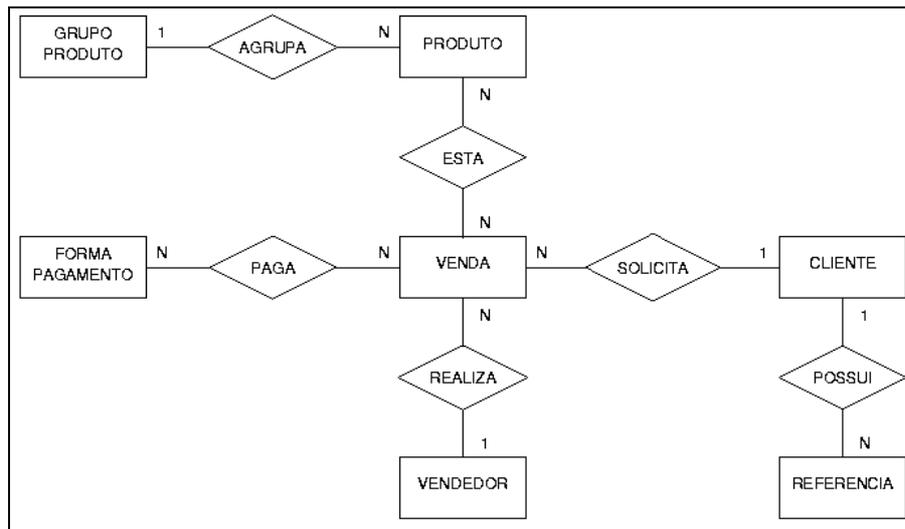


Figura 7 - Modelo Entidade-Relacionamento

Os modelos de dados podem ser basicamente de dois tipos:

- **Alto nível:** ou modelo de dados conceitual, que fornece uma visão mais próxima do modo como os usuários visualizam os dados realmente;
- **Baixo nível:** ou modelo de dados lógico ou físico, que fornece uma visão mais detalhada do modo como os dados estão realmente armazenados no computador.

f) Esquemas e Instâncias: Em qualquer modelo de dados utilizado, é importante distinguir a descrição do banco de dados do banco de dados por si próprio. A descrição de um banco de dados é chamada de **esquema de um banco de dados** e é especificada durante o projeto do banco de dados. Geralmente, poucas mudanças ocorrem no esquema do banco de dados. Os dados armazenados em um determinado instante do tempo formam um conjunto chamado de **instância do banco de dados**. A instância altera toda vez que uma alteração no banco de dados é feita.

O Sistema Gerenciador de Banco de Dados é responsável por garantir que toda instância do banco de dados satisfaça ao esquema do banco de dados, respeitando sua estrutura e suas restrições. O esquema de um banco de dados também pode ser chamado de **intensão** de um banco de dados e a instância de **extensão** de um banco de dados.

g) As Linguagens para Manipulação de Dados: Para a definição dos esquemas conceitual e interno pode-se utilizar uma linguagem chamada DDL (*Data Definition Language* -

Linguagem de Definição de Dados). O Sistema Gerenciador de Banco de Dados possui um compilador DDL que permite a execução das declarações para identificar as descrições dos esquemas e para armazená-las no catálogo do Sistema Gerenciador. A DDL é utilizada onde a separação entre os níveis interno e conceitual não é muito clara. Onde a separação entre os níveis conceitual e interno são bem claras, é utilizada outra linguagem, a SDL (*Storage Definition Language* - Linguagem de Definição de Armazenamento) para a especificação do esquema interno. A especificação do esquema conceitual fica por conta da DDL.

Em Sistemas Gerenciadores de Bancos de Dados que utilizam a arquitetura de três esquemas ou nível, é necessária a utilização de mais uma linguagem para a definição de visões, a VDL (*Vision Definition Language* - Linguagem de Definição de Visões).

Uma vez que o esquema esteja compilado e o banco de dados esteja populado, utiliza-se uma linguagem para fazer a manipulação dos dados, a DML (*Data Manipulation Language* - Linguagem de Manipulação de Dados).

2.3.3 Vantagens do uso de um Sistema Gerenciador de Banco de Dados

A utilização de Sistemas de Bancos de Dados tornou-se um padrão desde quando na década de 60 percebeu-se a inviabilidade de gerenciar um volume muito grande de informação sem um sistema de computador. A utilização de Sistemas Gerenciadores de Bancos de Dados trouxe várias vantagens, dentre elas:

a) Controle de Redundância: No processamento tradicional de arquivos, cada grupo de usuários deve manter seu próprio conjunto de arquivos e dados. Desta forma, acaba ocorrendo redundâncias que prejudicam o sistema com problemas como:

- Toda vez que for necessário atualizar um arquivo de um grupo, então todos os grupos devem ser atualizados para manter a integridade dos dados no ambiente como um todo;

- A redundância desnecessária de dados leva ao armazenamento excessivo de informações, ocupando espaço que poderia ser utilizado com outras informações.

b) Compartilhamento de Dados: Um Sistema Gerenciador de Banco de Dados multiusuário deve permitir que usuários distintos acessem o banco de dados ao mesmo tempo. Este fator é essencial para que múltiplas aplicações integradas possam manipular seus dados. Dessa forma é necessário manter o controle da concorrência para assegurar que o resultado de atualizações sejam corretos. Um banco de dados multiusuários deve fornecer recursos para a construção de múltiplas visões.

c) Restrição a Acesso não Autorizado: Um Sistema Gerenciador de Banco de Dados deve fornecer um subsistema de autorização e segurança, o qual é utilizado pelo DBA para criar contas e especificar as restrições destas contas. O controle de restrições se aplica tanto ao acesso aos dados quanto ao uso de *softwares* inerentes ao Sistema Gerenciador de Banco de Dados.

d) Representação de Relacionamentos Complexos entre Dados: Um banco de dados pode incluir uma variedade de dados que estão interrelacionados de várias formas. O Sistema Gerenciador de Banco de Dados deve fornecer recursos para se representar uma grande variedade de relacionamentos entre os dados, bem como recuperar e atualizar os dados de maneira prática e eficiente.

e) Tolerância a Falhas: Um Sistema Gerenciador de Banco de Dados deve fornecer recursos para recuperação de falhas tanto de software quanto de hardware.

Por fim e para melhor compreensão é interessante analisar e compreender a estrutura de um Sistema Gerenciador de Banco de Dados conforme figura 8 e 9.

Esta figura, de forma simplificada, mostra os componentes típicos de um SGBD. O banco de dados e o catálogo de dados são normalmente armazenados no disco. O acesso ao disco é controlado principalmente pelo sistema operacional, que organiza as entradas e as saídas. Um módulo de gerenciamento dos dados armazenados de alto nível do SGBD controla o acesso à informação do SGBD que está armazenada no disco, se for parte do banco de dados ou do catálogo.

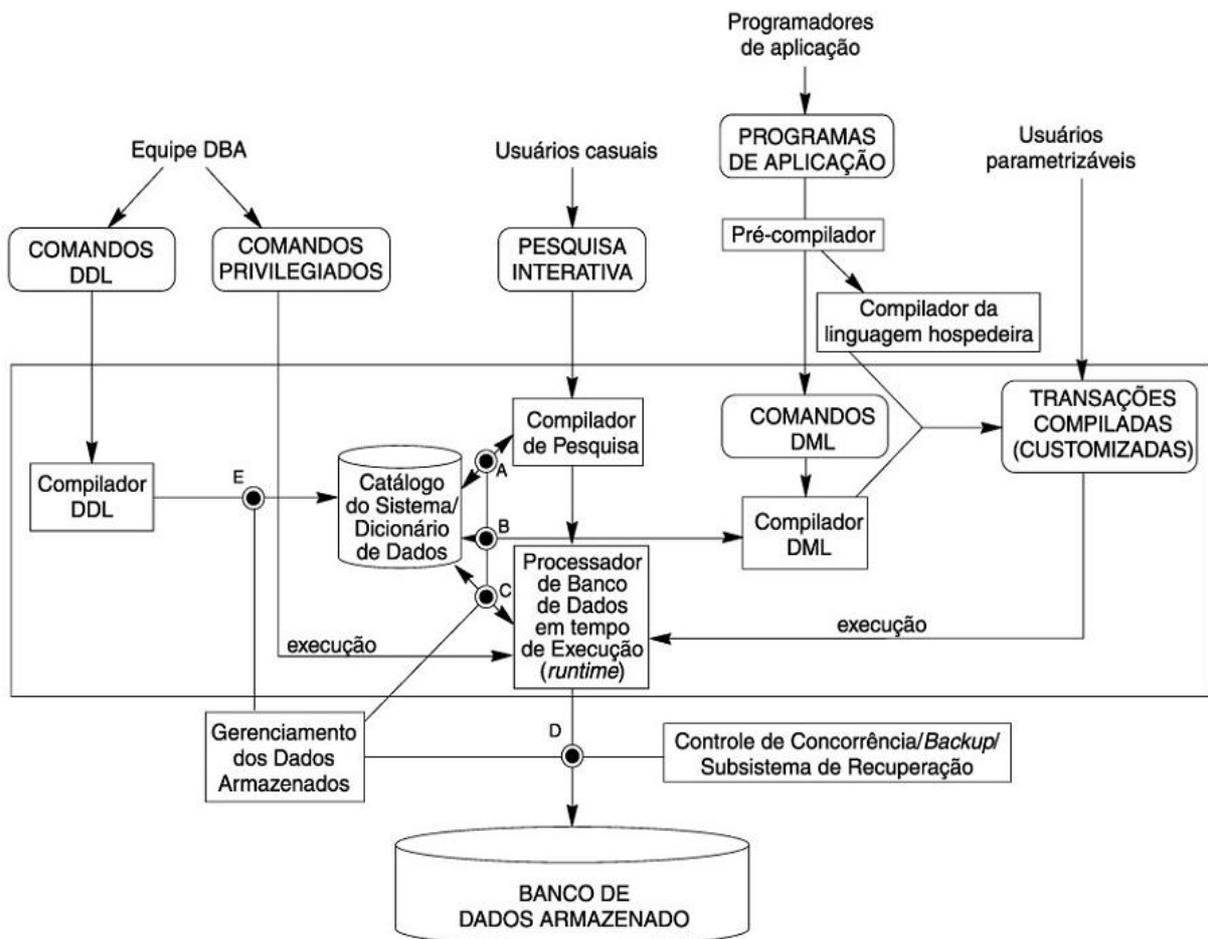


Figura 8 – Módulos de um SGBD

A figura 9, logo abaixo apresenta outra visão, desta vez da arquitetura interna de um SGBD centralizado.

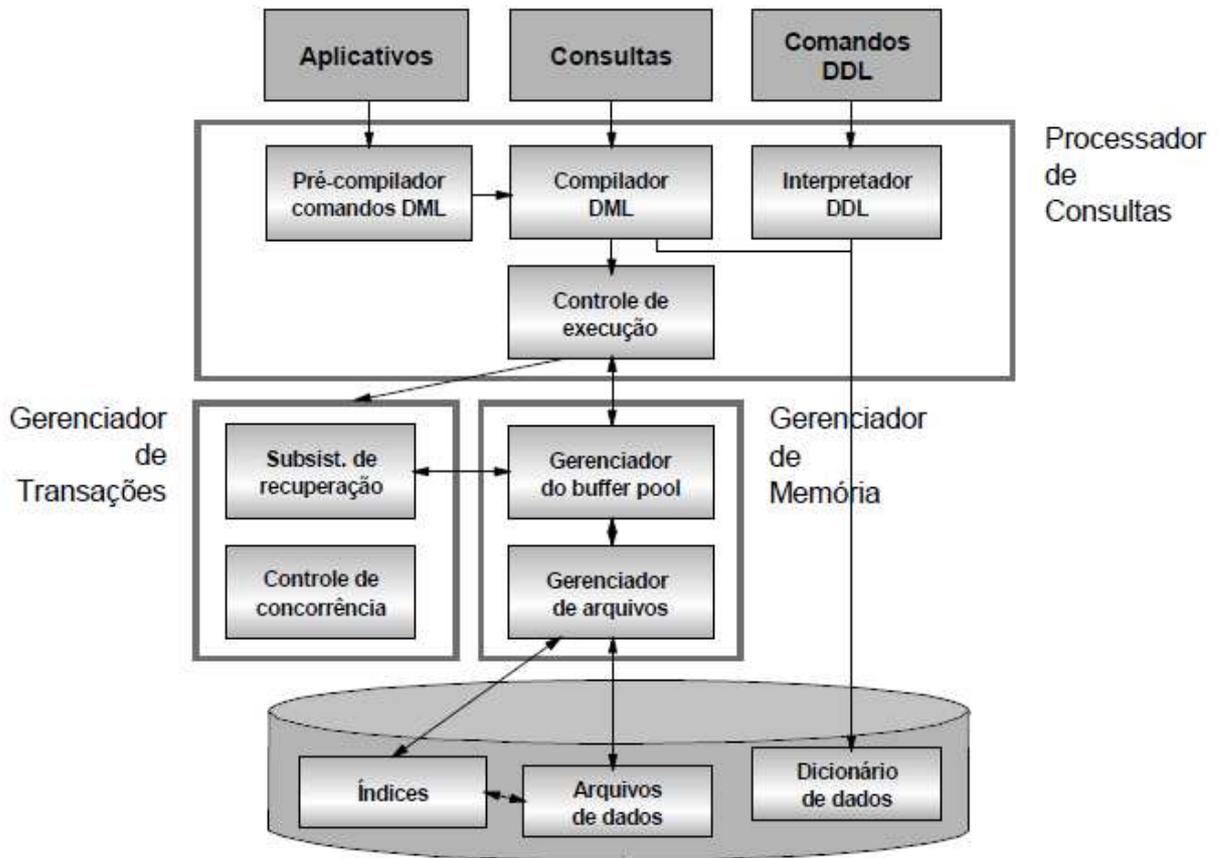


Figura 9 - Arquitetura de SGBD