INTRODUÇÃO À BANCO DE DADOS

O poder dos dados: atingir o público de maneira mais assertiva.

Tudo são dados: gps, buscadores, jogos, redes sociais...

Com o decorrer do tempo, supercomputadores foram aparecendo. A tecnologia evoluiu muito. A computação evolui em processamento.

Hoje em dia não temos só texto e número, temos redes sociais (com estruturas complexas: imagens, vídeos, likes), o cenário mudou!

3 V's: Velocidade / Variedade / Volume - BIG DATA

O que são Banco de Dados?

Dados são fatos e Banco de dados é um conjunto de dados que transmitem informação

- Tamanho - Complexidade

Banco de dados não é a mesma coisa que SGBD.

Banco não é tão coerente como uma SGBD

SISTEMAS DE GERENCIAMENTO DE BANCO DE DADOS

SGBD: facilitam a nossa vida, tem restrição de acesso. É composto por diversos banco de dados.

Etapas: Definição, Construção, Manipulação, Compartilhamento

Conseguimos executar uma série de ações e submetemos a ação ao SGBD e ele retorna a informação. As Query's são solicitações.

Cada atributo leva um Tipo de Dado: INT, VARCHAR...

Ideia de SGBD surgiu em 1960 - através da IBM com o intuito de diminuir custos com pessoal. E torna a vida do desenvolvedor mais fácil.

RELACIONAL 1970 - criado por Edgar (Ted) Codd - pela IBM

1980 - democratização de acesso a tecnologia aumentou

1990 - programação voltada a orientação a objeto - POO

2000 - surgiram os NoSQL, para suprir um gap que o modelo relacional deixou

Modelo Hierárquico: voltado para a informação, estamos buscando registros, usando instância e não tabela. É uma busca em árvore, complicada e menos otimizada

Modelo em Rede: É estrutura. Relacionamento N:M, ele não é tão rico quando o relacional. A pessoa que lida com esse modelo precisa conhecer a estrutura física com BD.

Modelo Relacional: você tem estrutura, é um modelo mais rico. Você otimiza as consultas e tem transparência no sistema

- LDD ou DDL: são comandos específicos para a definição dos dados Ex: MySQL, MariaDB

SGBD mais utilizados no mercado (não há escolha errada, cada um tem uma característica)

- 1 Oracle: ele foi o pioneiro, é usado por grandes corporações.]
- 2 MySQL: mantido pela Oracle, é de código aberto. É muito leve e utilizado em aplicações web. Boa integração com PHP.
- 3 SQLServer: mantido pela Microsoft, ferramentas para integração com BI.
- 4 PostgreSQL: considerado exemplo quando se fala em especificação, ele é de 1995. Bem robusto, open source, você consegue armazenar documentos (JSON, por ex).
- 5 mongoDB: orientado a documentos. Banco que armazena dados em blocos. Ótimo para quando não se tem muita previsibilidade (ex: blackfriday)
- 6 redis: diminui e evita atrasos de busca, ele pega em memória, roda direto na RAM.
- 7 ElasticSearch: mecanismo de busca distribuído. Pode descobrir algo que não é esperado. Interface simples, utiliza JSON
- 9 Access: integra com office. Mais simples
- 11 cassandra: gratuito, open source, NoSQL, desenvolvido pelo Facebook.
- 12 MariaDB: originou do MySQL, por isso é bem similar. Método de busca e consulta mais otimizado.

ERA DOS DADOS E FUTURO DA MODELAGEM

Contexto de dados

4º Paradigma (paradigma científico)

Paralelismo: não há uma sequência entre as ações, associado a capacidade de processar uma determinada aplicação de maneira paralela. Conjunto de operações que podem ser executadas em paralelo.

Big Data: também é paralelo mas diferente, você persiste no dado

Cloud: AWS, você utiliza banco de dados como serviço. Você não tem preocupações com falhas, por exemplo.

HPC: computação de alto desempenho. Não persiste, simplesmente processa.

BigData: processa e armazena

MERCADO DE DATA Engenheiro de dados Cientista de dados Analista de dados

Data-driven

A partir de uma análise eu faço uma interpretação. O consumidor é o centro do negócio.

Modelos NoSQL

Surgiram pra preencher o gap dos bancos relacionais.

Distribuído, tem alta performance e modelagem diferenciada.

Utilizado em cenários pontuais. NOT ONLY SQL

Documentos - Wide-columns - Key Value - Grafos - Orientado à objetos