



INSTITUTO FEDERAL
Paraíba
Campus João Pessoa

Aula

11

Banco de Dados II

Atualizada em 18/02/2019

BD No-SQL Uma visão geral

Professor:

Dr. Alex Sandro da Cunha Rêgo



alex@ifpb.edu.br



- **Motivação**

*“A partir do surgimento da web, o volume de dados armazenados sobre usuários, objetos, produtos, etc. **tem crescido exponencialmente**. Nesta direção, os dados também passaram a ser acessados com maior frequência e processado de forma mais intensiva”*

- Exemplo: **Redes Sociais**

- RDBMS não foram projetados para lidar com os desafios de **escalabilidade** e **agilidade** que as aplicações modernas requerem

- ✓ Aproveitamento do **baixo custo de armazenamento** e **poder de processamento** disponível na “nuvem”

- **NoSQL**

- Non-relational and not SQL



- **BDs SQL Relacionais**

- ❑ Bancos de dados estruturados sob a forma de linhas e colunas, armazenadas em tabelas e relacionadas entre si
- ❑ Banco de Dados predominante no mercado
- ❑ SQL é a linguagem padrão
 - ✓ **DML** (*Data Manipulation Language*): SELECT, INSERT, UPDATE e DELETE
 - ✓ **DDL** (*Data Definition Language*): CREATE, ALTER e DROP
 - ✓ **DCL** (*Data Control Language*): GRANT e REVOKE
 - ✓ **DTL** (*Data Transaction Language*): COMMIT e ROLLBACK
- ❑ **Soluções:** SQL Server, Oracle, Informix, MySQL, PostgreSQL



- **BDs No-SQL**

- Banco de Dados “Orientado a Documentos”
 - ✓ Bancos de dados que não utilizam o formato padrão de entidade x relacionamento
- A base de dados é constituída por **coleções** de documentos em formato semelhante ao JSON
 - ✓ As coleções seriam algo como um **database**
 - ✓ Os documentos são listas de **propriedade-valor** ou **array** ou **documentos aninhados**
- Linguagem de Consulta/Manipulação de dados
 - ✓ Utiliza uma chamada de métodos semelhante a javascript, com argumentos semelhantes à estrutura de JSON

Introdução



- **BD SQL Relacional vs BD No-SQL**

ISBN	title	author	format	price
9780992461225	JavaScript: Novice to Ninja	Darren Jones	ebook	29.00
9780994182654	Jump Start Git	Shaumik Daityari	ebook	29.00

registros



Estrutura rígida; as colunas possuem tipos pré-definidos e não aceitam informações de um dado incompatível

documento

```
{  
  name: "sue",  
  age: 26,  
  status: "A",  
  groups: [ "news", "sports" ]  
}
```

Pares **campo-valor**
ou **array**



Documentos similares podem ser armazenados em uma coleção. Qualquer tipo de dado pode ser armazenado em um documento

Esquema de Dados



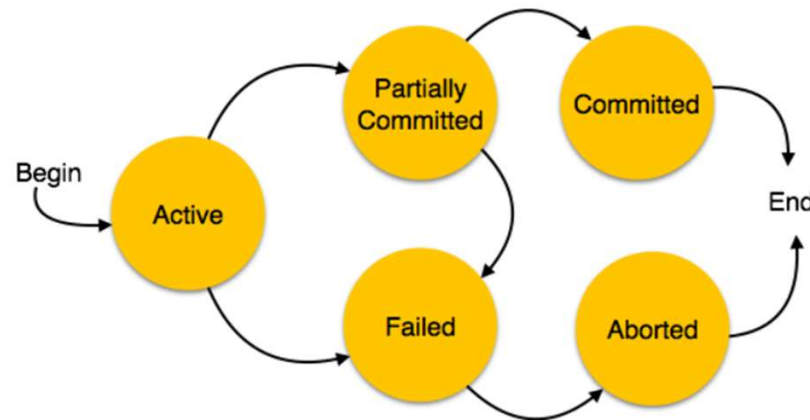
- BD SQL
 - Para **inserir dados**, primeiramente deve ser criada a tabela e definidas as colunas e tipos de dados
 - ✓ A **definição do esquema** de dados é o primeiro passo
 - ✓ Mudanças profundas no esquema podem complicar...
 - O esquema contém informações a respeito de **chaves primárias, índices, relacionamentos e funcionalidades** tais como **triggers e procedures**
- BD No-SQL
 - Dados podem ser inseridos em **qualquer lugar** e em **qualquer momento**
 - ✓ Os campos e tipos são definidos no momento da inserção na coleção
 - Coleções não são “normalizadas”

Integridade de Dados



- BD SQL
 - Integridade dos dados é garantida utilizando **constraints** (chave estrangeira, chave primária, remoção em cascata, etc.) **procedures**
- BD No-SQL
 - Regras de integridade não estão disponíveis em BDs No-SQL
 - ✓ O documento **deverá ser autosuficiente** em relação às informações sobre um item

Controle de Transações



- BD SQL
 - ❑ A execução de instruções de atualização/inserção em sequência é garantida com o uso de **transações** (tudo ou nada)
- BD No-SQL
 - ❑ Modificações em um **único** documento são **atômicas** por natureza
 - ❑ **Não há** controle transacional equivalente para update em **múltiplos documentos**

CRUD: Sintaxe



- BD SQL
 - ❑ Utiliza uma linguagem declarativa (DML) que tem se tornado um padrão internacional
- BD No-SQL
 - ❑ Utiliza métodos similares a javascript com argumentos semelhantes a JSON
 - ✓ Operações básicas são simples, porém, JSON aninhados podem se tornar confusas para **queries mais complexas**

SQL	No-SQL
<pre>UPDATE book SET price = 19.99 WHERE ISBN = '9780992461256'</pre>	<pre>db.book.update({ ISBN: '9780992461256' }, { \$set: { price: 19.99 } });</pre>

Outras Características



- **Performance**

- **BD SQL**

- ✓ Necessita de **junção de tabelas** e consultas SQL mais complexas

- **BD No-SQL**

- ✓ Todas as informações sobre um item específico **estão em um único lugar** ('desnormalização').

- **Escalabilidade**

- **BD SQL**

- ✓ Clusterização (múltiplos servidores) é uma possibilidade para o crescimento de dados

- **BD No-SQL**

- ✓ Devido à simplicidade do modelo de dados, muitos BD's No-SQL tem sido desenvolvidos a contemplar este item

Vantagens e Desvantagens



- **BD No-SQL**

A FAVOR	CONTRA
Ausência de um esquema (BD vai sendo criado conforme forem sendo inseridos e atualizados documentos. Sem tabelas, esquemas, linhas ou SQL)	Redundância (A mesma informação pode ser replicada em diferentes documentos – Sem foco no relacionamento entre documentos)
Facilidade de uso (Curva de aprendizado é baixa para começar a se trabalhar com um BD No-SQL. Não necessita de se fazer configurações complicadas ou instalações complexas)	Ausência de Joins (Não se tem a mesma performance quando se necessita linkar diferentes documentos)
Desempenho nas consultas (Um documento normalmente é autossuficiente de informações – sem Joins)	Ausência de Transações (Não garante a atomicidade em grandes operações de manipulação de dados – Apenas em um único documento)



- **Como as empresas estão usando NoSQL**
 - ❑ **Facebook Messaging Platform:** indexação de mensagens e termos, para orientar as pesquisas por todo o conteúdo das mensagens (Apache Cassandra)
 - ❑ **BBC iPlayer Online Media Catalog:** serviço online para prover acesso a metadados de programas de rádio e televisão
 - ❑ **BBC Dynamic Semantic Publishing (DSP):** Sugestão de tópicos para matérias, aos jornalistas
 - ❑ **Secure Information Sharing:** acesso à informação sem descuidar da segurança do dado. Por exemplo, ler partes de um livro, para apreciação, sem conceder acesso *full* antes de comprá-lo

Introdução



- **Como as empresas estão usando NoSQL**
 - **New York Times:** projeto que lida com coleta, filtragem e análise de dados em tempo real a partir de diferentes fontes de dados. A ideia é identificar como os leitores interagem com as notícias e entre si
 - ✓ **Exemplo:** Postagem de blogs, comentários, entre outros



- **Cenários de USO**

- **Big Data**

- ✓ Grande volume de informações, mineração de dados

- **Escrita intensa**

- ✓ Implementa uma solução assíncrona e performática para lidar com escritas contínuas

- **Busca de dados simples**

- ✓ Sem cruzar informações de diferentes coleções

- **Alta escalabilidade e disponibilidade**

- ✓ Permite escalonamento vertical e horizontal

- **Schema de dados é instável**

- ✓ Poucas “colunas” se mantêm preservadas

Introdução



- No-SQL não é uma substituição ao SQL
 - Ambos possuem diferentes abordagens
 - Alguns projetos requerem um BD SQL, enquanto outros saem melhor com um BD No-SQL
 - BDs No-SQL **não tenta ser o melhor em tudo e admite prontamente que não é ideal para todos os cenários.**
 - ✓ Um banco de dados otimizado quanto à questão de **simplicidade, velocidade e escalabilidade**, porém, abrindo mão das propriedades ACID (garantia de consistência e armazenamento de dados)

Referências Bibliográficas



- What is NoSQL?. Disponível em <https://www.mongodb.com/nosql-inline>
- Cassandra - A Decentralized Structured Storage System. Disponível em <https://docs.datastax.com/en/articles/cassandra/cassandra-athenandnow.html>
- What is NoSQL?. Disponível em <https://www.mongodb.com/nosql-inline>