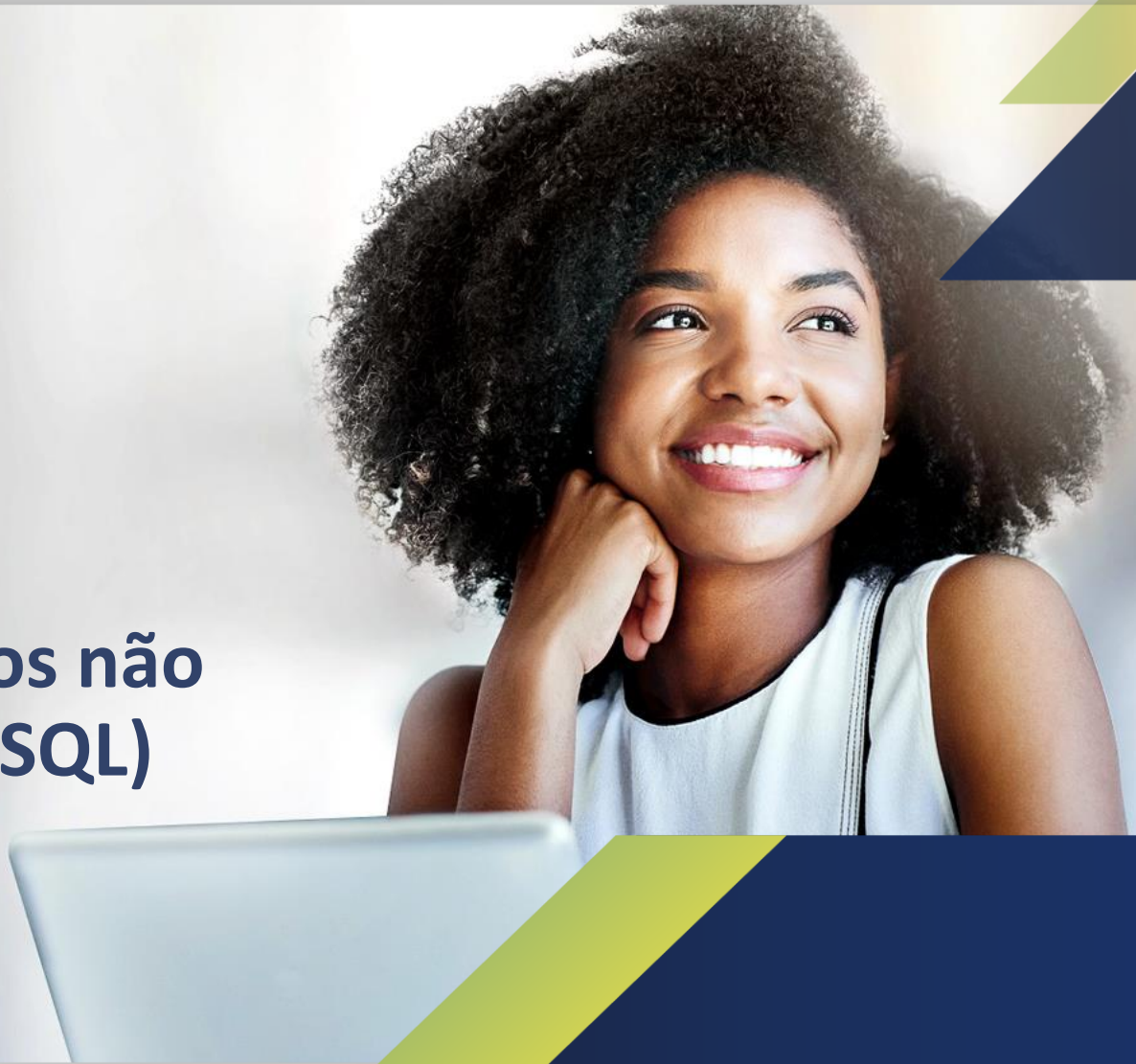


PÓS-GRADUAÇÃO

Banco de dados não relacional (NoSQL)



PÓS-GRADUAÇÃO

Modelo orientado a família de colunas

Bloco 1

Marcelo Tavares de Lima





► Objetivos

- Identificar as características de bancos de dados orientados a família de colunas.
- Compreender a estrutura de um modelo de banco de dados orientado a família de colunas.
- Compreender a diferença entre este modelo e os demais modelos adotados pelo padrão NoSQL.
- Conhecer as vantagens e desvantagens em adotar este modelo de estrutura de banco de dados NoSQL.




► Modelo orientado a família de colunas

- Modelo NoSQL colunar.
- Persistência de dados.
- Persistência de dados refere-se ao armazenamento não volátil da informação.
- Não volátil se refere ao armazenamento na íntegra, de maneira permanente.
- Persistência está atrelada à rápida resposta do banco (estrutura otimizada).



► Modelo orientado a família de colunas

- Tem como uma das principais características, a maior velocidade de resposta na realização de consultas quando comparado aos modelos relacionais.
 - Outras características importantes: segurança, não redundância, consistência e flexibilidade do banco.
- 




► Modelo orientado a família de colunas

- No entanto, a velocidade na consulta é um dos principais fatores que fazem com que grandes empresas, como Facebook, Netflix, dentre outros, utilizem modelos não relacionais para armazenar seus dados.



► Modelo orientado a família de colunas

- Armazenamento em bancos relacionais atrela registros de dados a uma linha dentro de uma tabela, permitindo cruzamento com informações contidas em outra tabela com comando JOIN.
 - Armazenamento em bancos não relacionais, especificamente, em modelo colunar, organiza dados em colunas independentes e não existem relacionamentos.
- 

► Modelo orientado a família de colunas

**Figura 1 - Armazenamento em
banco relacional**

id	nome	email
20	José Miguel	josemiguel@dominio.com
21	Antônio dos Santos	antoniodossantos@dominio.com
22	Pedro Lucas	pedrolucas@dominio.com

Fonte: elaborado pelo autor.

► Modelo orientado a família de colunas

Figura 2 - Armazenamento em
modelo não relacional

ID

id	
id	valor
1	20
2	21
3	22

NOME

nome	
id	valor
1	José Miguel
2	Antônio dos Santos
3	Pedro Lucas


EMAIL

email	
id	valor
1	josemiguel@dominio.com
2	antoniiodossantos@dominio.com
3	pedrolucas@dominio.com

Fonte: elaborado pelo autor.



► Modelo orientado a família de colunas

- Supercolunas: são colunas que, em sua estrutura, possuem outras colunas.
 - Permitem o armazenamento de muitos dados dentro da mesma coluna, otimizando os resultados de busca.
 - Uma coleção de supercolunas gera uma estrutura chamada de superfamílias de colunas.
- 

► Modelo orientado a família de colunas

Figura 3 - Supercoluna em um banco não relacional

chave	valor	
18	endereco	
	tipo	Rua
	end	Andorinhas
	numero	50
	bairro	Centro
	municipio	Santos
	uf	sp

Fonte: elaborado pelo autor.




► Modelo orientado a família de colunas

- Keyspace: segundo Secco *et al.* (2016, p. 44), “agrupamento de famílias de colunas (similar a um banco de dados do modelo relacional)”.
- Keyspace: funciona como dicionário de dados, aponta para todas as colunas, famílias de colunas, supercolunas, superfamília de colunas e respectivas ligações entre si.



► Modelo orientado a família de colunas

- Uma das vantagens é seu poder de escalabilidade, ou seja, sua capacidade de manipulação de dados de maneira uniforme à medida que estes crescem.
 - O poder de adaptação a diferentes cenários de grandes massas de dados é outra característica importante.
- 



► Modelo orientado a família de colunas

- Uma desvantagem é que quanto mais complexa for a estrutura de suas colunas, maior será a complexidade de manipulação.

PÓS-GRADUAÇÃO

Modelo orientado a família de colunas

Bloco 2

Marcelo Tavares de Lima






► Banco de dados Cassandra

Banco de dados distribuído massivamente escalável, criado para armazenar uma grande quantidade de dados espalhados por vários servidores e, ainda assim, oferecer alta viabilidade de acesso e dados consistentes, segundo Lakshman, Malik *apud* Anicetoxavier (2014, p. 23).




► Banco de dados Cassandra

- Criado e desenvolvido, em 2008, como um projeto *open source* pelo Facebook, com a intenção de melhorar seu mecanismo de pesquisa.
 - No ano de 2009, foi adotado pela Apache Software Foundation.
 - Utilizado pelo Netflix, Ebay, NASA, órgãos governamentais de diversos países etc.
- 



► Banco de dados Cassandra

- Seu desenvolvimento tem arquitetura de sistema distribuído, baseada em arquitetura do Dynamo da Amazon.
 - Tem modelo de dados baseado no BigTable, da Google.
 - Também permite consulta como modelo chave/valor, por isso, é considerado como híbrido.
- 



► Banco de dados Cassandra

- Principais características:
 - Distribuído.
 - Descentralizado.
 - Escalável.
 - Alta disponibilidade.
 - Tolerante a falhas.
 - Consistente.
 - Alta performance.



► Banco de dados Cassandra

- Utiliza a linguagem *Cassandra Query Language (CQL)*, desenvolvida para o próprio banco.
- Assemelha-se à linguagem *Structured Query Language (SQL)*.

► Banco de dados Cassandra

Figura 4 - Script CQL

```
CREATE TABLE cadastro(  
  id INT,  
  nome TEXT,  
  email TEXT,  
  PRIMARY KEY(id)  
);
```

Fonte: elaborado pelo autor.

► Banco de dados Cassandra

Figura 5 - Consulta CQL

```
SELECT * | select_expression | DISTINCT partition  
FROM [keyspace_name.] table_name  
[WHERE partition_value  
    [AND clustering_filters  
    [AND static_filters]]]  
[ORDER BY PK_column_name ASC|DESC]  
[LIMIT N]  
[ALLOW FILTERING]
```

Fonte: elaborado pelo autor.

PÓS-GRADUAÇÃO

Teoria em prática

Bloco 3

Marcelo Tavares de Lima





► Aplicação do modelo orientado a família de colunas

- Suponha que você seja um *Data Base Administrator (DBA)* e precisou utilizar uma estrutura NoSQL para armazenar os dados de uma rede social.
- Você necessitou criar uma estrutura para armazenar os dados que trafegassem pela rede: nome dos usuários, e-mails, postagens, fotos e vídeos.
- Elabore a estrutura desse banco não relacional baseado em família de colunas, utilizando diagramas para demonstrar a forma como os dados seriam organizados.

► Aplicação do modelo orientado a família de colunas

Figura 6 - Modelo orientado a família de colunas

Keys	Valor	
1	Dados	
	Dados	
	Nome	
	e-mail	
	Postagem	
	Videos	
2	Dados	
	Nome	
	e-mail	
	Postagem	
	Videos	
3	Dados	
	Nome	
	e-mail	
	Postagem	
	Videos	
...		

Fonte: elaborado pelo autor.

Dica do professor

Bloco 4

Marcelo Tavares de Lima





► Artigo para leitura

- Apache Cassandra.
- Kerlla de Souza Luz, Thiago Vinícius de Melo Almeida.
- Departamento de Ciência da Computação – Universidade de Brasília (UNB).
- Disponível em:
<<https://cic.unb.br/~alchieri/disciplinas/posgraduacao/sd/artigoG2.pdf>>. Acesso em: 24 out. 2019.

► Referências

ANICETO, R.C.; XAVIER, R.F. **Um estudo sobre a utilização do banco de dados NoSQL Cassandra em dados biológicos**. 2014. Monografia (Graduação em Ciência da Computação) – Instituto de Ciências Exatas, Universidade de Brasília, Brasília, 2014. Disponível em: <http://bdm.unb.br/bitstream/10483/7927/1/2014_RodrigoCardosoAniceto_ReneFreireXavier.pdf>. Acesso em: 24 out. 2019.

► Referências

SECCO, R.R.; SILVA, F.A.; MARACCI, F.V.; PAZOTI, M.A.
Análise comparativa entre o banco de dados Cassandra (modelo NoSQL) e o PostgreSQL (modelo relacional) em duas diferentes organizações empresariais. **Colloquium Exactarum**, v. 8, n.2 , 2016, p.39 - 56. Disponível em:
<<https://webcache.googleusercontent.com/search?q=cac he:vN1pS-09OVwJ:https://revistas.unoeste.br/index.php/ce/article/download/1324/1754/+&cd=13&hl=pt-BR&ct=clnk&gl=br>>. Acesso em: 24 out. 2019.

