

NoSQL: necessidades e tipos

Marino H. Catarino
marinohc@gmail.com

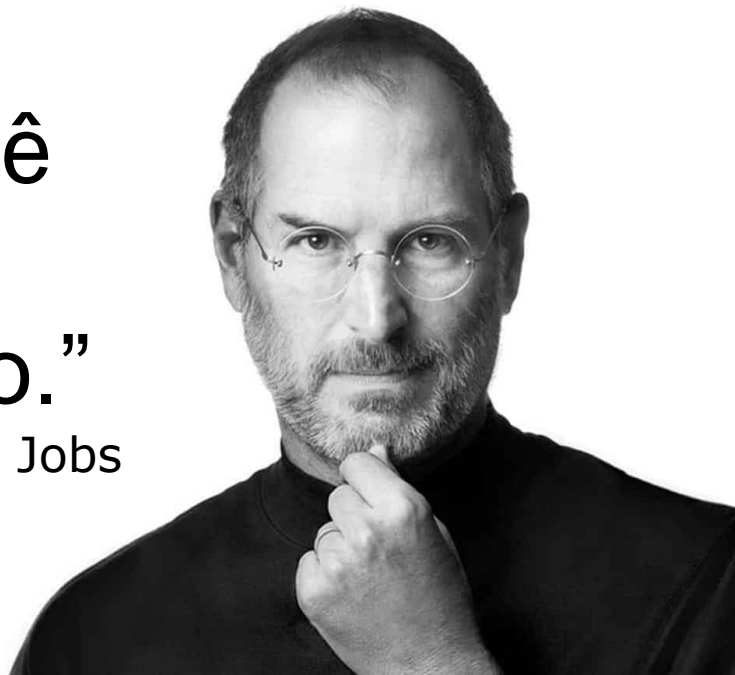
Abril/2021





“Tenha coragem de seguir o
que seu coração e sua
intuição dizem.
Eles já sabem o que você
realmente deseja.
Todo o resto é secundário.”

Steve Jobs





Problema:

Banco de dados





Armazenamento relacional:

ORACLE®



Microsoft®
SQL Server®



PostgreSQL





Problema:

Banco de dados





Motivos: Redes sociais





Motivos: sensores





Questão:

Onde entra o Armazenamento de Dados?





Solução:

Sistemas Distribuídos + Bancos de Dados



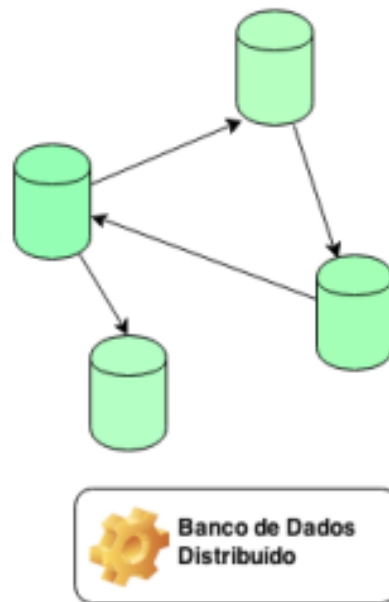
O que é um banco NoSQL?

Um banco de dados distribuído, não relacional, criado para trabalhar com grandes quantidade de dados (big data) utilizando servidores auxiliares.



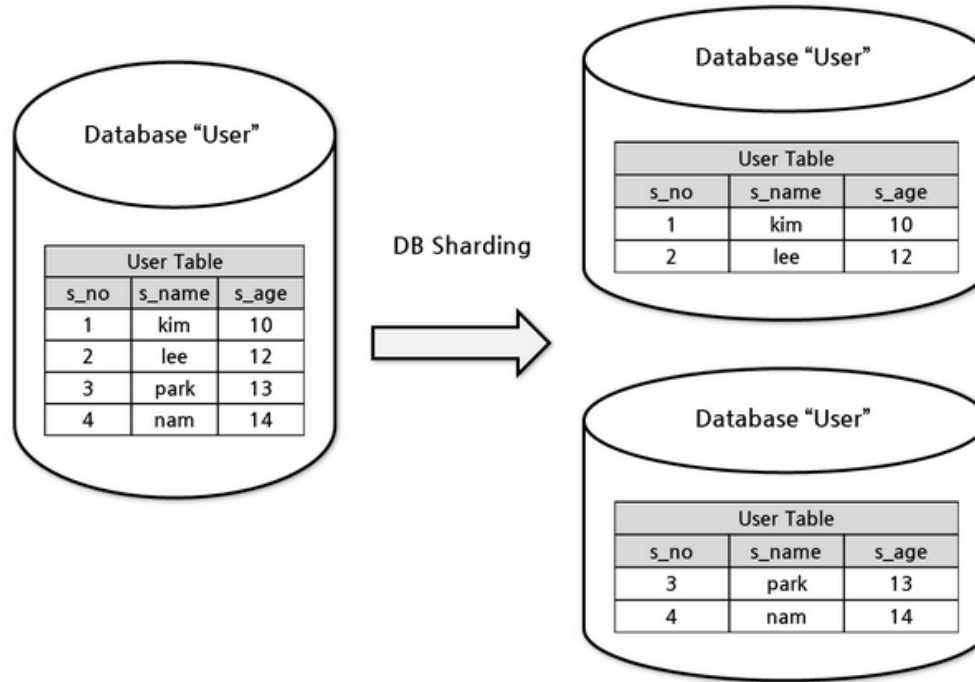
Distribuído?

- Particionamento de Dados
- Balanceamento de carga
- Tolerância a Falhas





Particionamento horizontal





Solução: NoSQL



Cassandra



mongoDB



membase





Características dos NoSQL

- Alta escalabilidade
- Tolerância a falhas
- Alta vazão (throughput)
- Arquitetura Distribuída (cluster ou cloud)
- Open Source (a maioria dos sistemas NoSQL atuais)



Origem dos NoSQL

Sistemas internos desenvolvidos pela Google (Bigtable), e Amazon (Dynamo)

A arquitetura destes dois sistemas influencia a grande maioria dos sistemas NoSQL até hoje.



Tipo de armazenamento

Existem diversos tipos de armazenamento, no qual cada um trata os dados de uma forma diferente e que pode ser mais específico para o objetivo desejado. Os tipo de armazenamento são:

- Orientado a colunas
- Chave valor
- Documentos
- Banco de dados em Grafos



Orientado a Colunas

Fortemente inspirados pelo BigTable do Google, suportam várias linhas e colunas, e também permitem subcolunas. Além do BigTable do Google, outros que usam essa tecnologia são:

- HBase(Apache)
- HiperTable
- Cassandra(Apache)



Orientado a colunas

Colunar

Red					Purple				
	Yellow								
		Yellow						Pink	
			Green						
			Blue			Dark Blue			Orange
			Brown						
							Cyan		
					Pink				

Tabela			
Família de coluna 1		Família de coluna 2	Família de coluna 3
Coluna 1	Coluna 2	Coluna 3	Coluna 4
#1 {Chave: Valor, Chave: Valor}	#1 {Chave: Valor, Chave: Valor}	#1 {Chave: Valor, Chave: Valor}	#1 {Chave: Valor, Chave: Valor}
#2 {Chave: Valor, Chave: Valor}	#2 {Chave: Valor, Chave: Valor}	#2 {Chave: Valor, Chave: Valor}	#2 {Chave: Valor, Chave: Valor}



Chave Valor

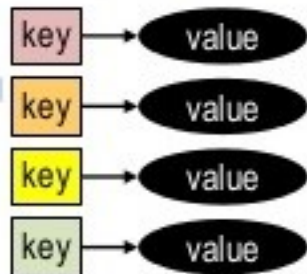
Esse é o tipo de banco de dados NoSQL mais simples. O conceito dele é uma chave e um valor para essa chave, mas ele é o que aguenta mais carga de dados.

Estes tipos de bancos de dados são o que tem a maior escalabilidade:

- Dynamo DB
- Berkeley DB
- Tokyo Cabinet
- Project Voldermort
- MemcacheDB
- SimpleBD
- Redis
- Riak



Chave/Valor



Chave	Valor
Nome	Marino
Sobrenome	Catarino
Gênero	Masculino
CPF	123.456.789-23



Documentos

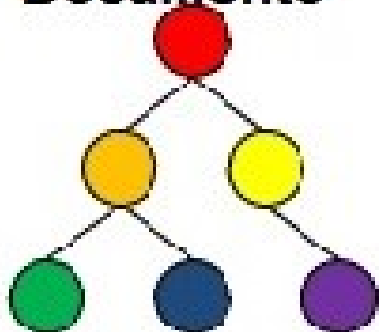
Baseados em documentos XML ou JSON, podem ser localizado pelo seu id único ou por qualquer registro que tenham no documento:

- CouchDB(Apache)
- MongoDB
- RavenDB



Orientado a documentos

Documento



<Key=CustomerID>

```
{
  "customerid": "fc986e48ca6"
  "customer":
  {
    "firstname": "Pramod",
    "lastname": "Sadalage",
    "company": "ThoughtWorks",
    "likes": [ "Biking", "Photography" ]
  }
  "billingaddress":
  {
    "state": "AK",
    "city": "DILLINGHAM",
    "type": "R"
  }
}
```

Key



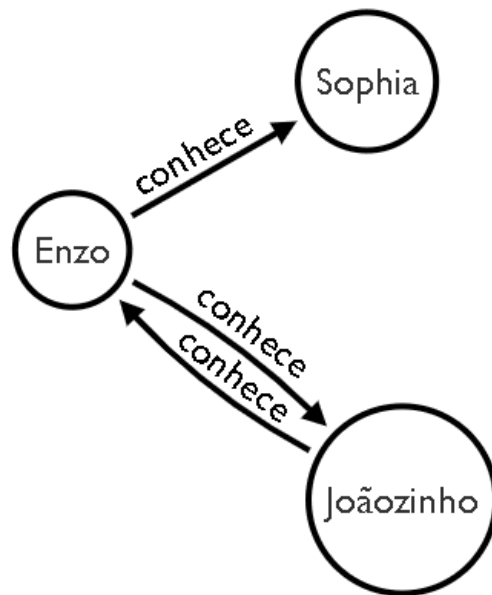
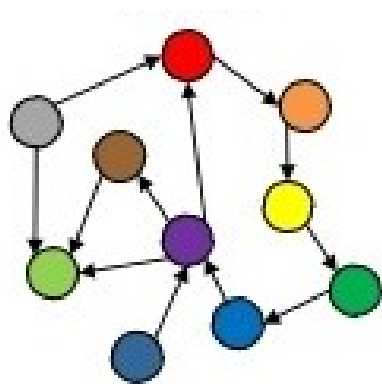
Banco de dados em Grafos

Com uma complexibilidade maior, esses bancos de dados guardam objetos, e não registros, como os outros tipos de NoSQL. A busca desses itens é feita pela navegação destes objetos:

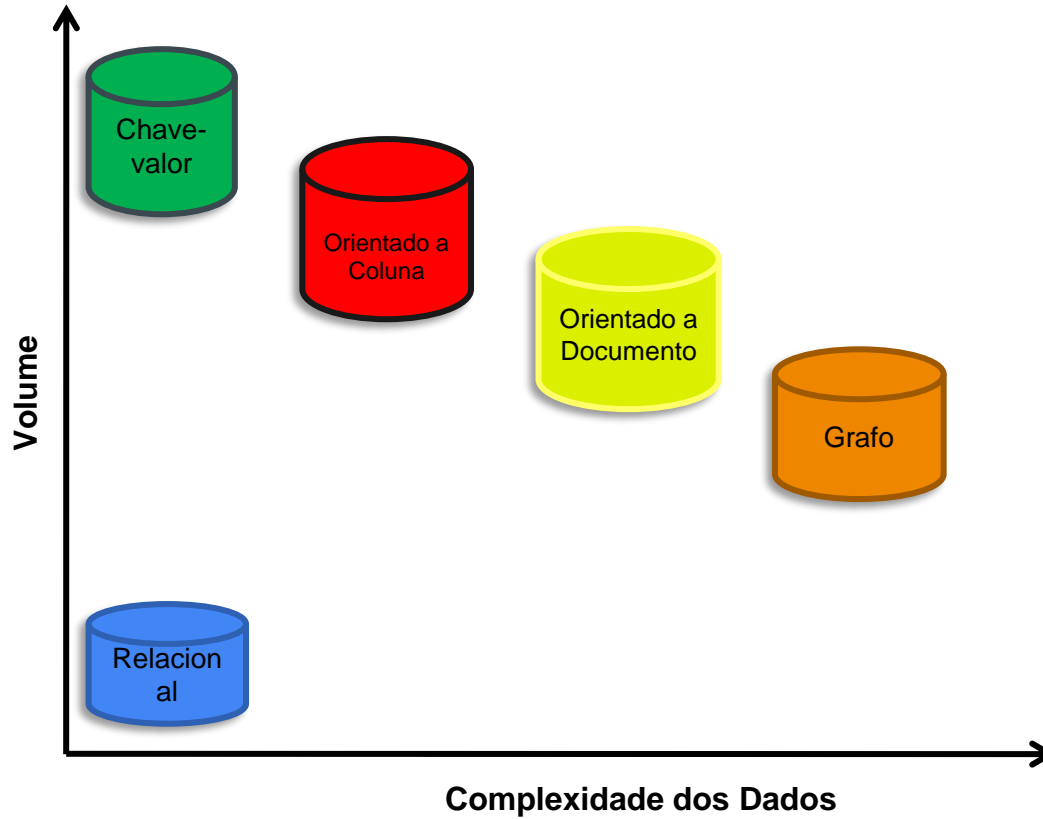
- Neo4J
- InfoGrid
- HyperGraphDB

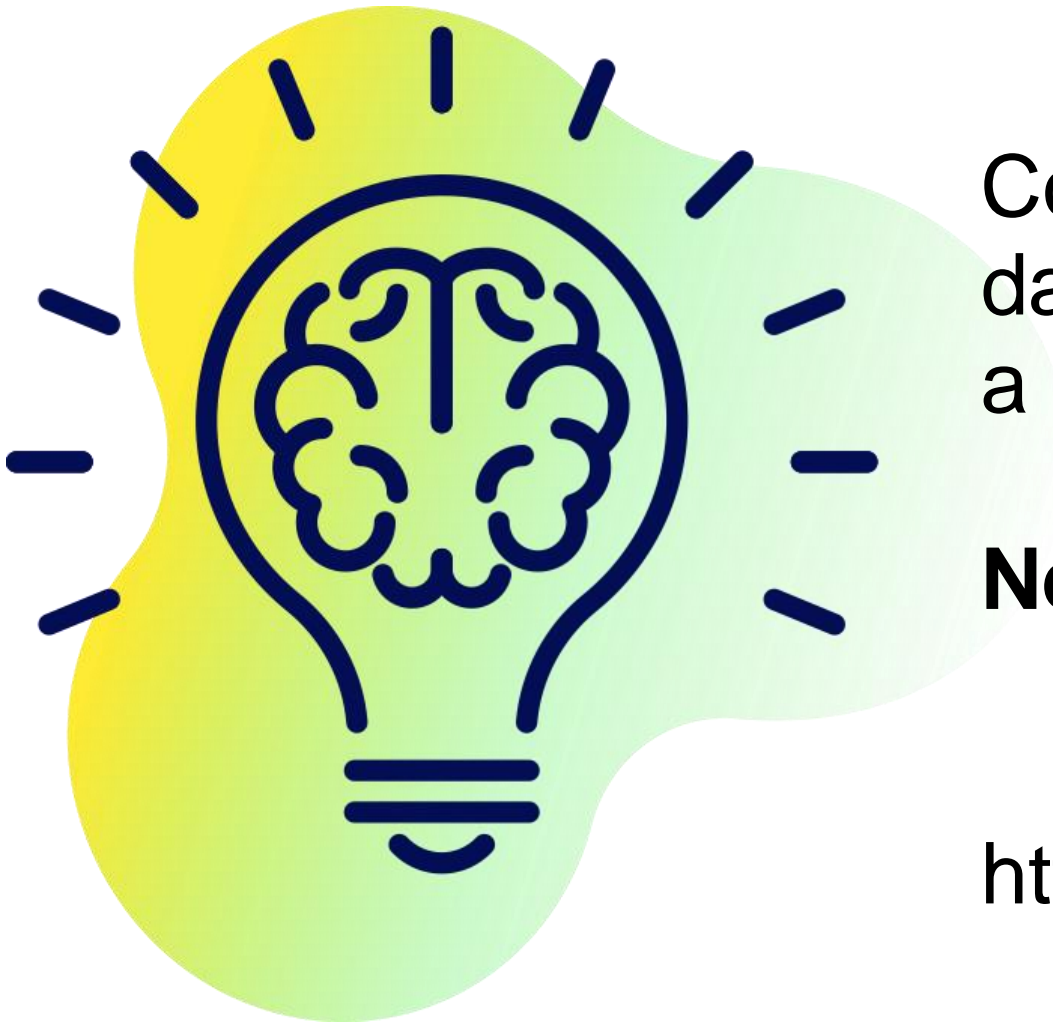


Grafos



Comparativo





Conhecer o banco de
dados orientado
a grafos na prática:

Neo4J

<http://neo4j.com>