



SGBDs NoSQL

Modelos e Sistemas de Colunas e Grafos

Aluno: Jorge Dias
Disciplina: IN940



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO





Roteiro

- Introdução
- Motivação
- NoSQL (Características)
- Comparativo de SGBDs NoSQL x SGBDs Relacionais
- Tipos de modelagens de dados NoSQL
- Modelagem de Colunas
- Modelagem de Grafos
- Os SGBDs Que as empresas Usam
- Conclusões
- Referências
- Espaço para perguntas



Introdução

- Os Bancos de dados relacionais começaram a ser apresentados no começo da década de 70 por Ted Codd, porém, só foram implementados com sucesso no final da década de 80.
- Mas a quantidade de dados que eram armazenados lá no começo dos anos 70, não são as mesmas quantidades de dados que são armazenados atualmente e as ferramentas utilizadas hoje evoluíram muito em relação as ferramentas daquela época.
- **Exemplos de tecnologias que temos hoje:**
 - Redes sociais
 - Computação nas Nuvens
 - Web 2.0



Motivação

- Quantidade de dados em escala gigantesca (Terabytes / Petabytes)
- Os Bancos de Dados Relacionais suportariam toda essa quantidade de dados? O processo seria bem otimizado?
- BDs relacionais trabalham com ACID (**A**tomicidade, **C**onsistência, **I**solabilidade, **D**urabilidade)
- BDs NoSQL trabalham com BASE (**B**asicamente Disponível, **E**stado Leve, **E**ventualmente consistente)



ACID x BASE

ACID	BASE
Consistência Forte	Consistência Fraca
Isolamento	Foco em Disponibilidade
Concentra-se em Commit	Melhor Esforço
Transações Aninhadas	Respostas Aproximadas
Disponibilidade	Mais Simples e Mais Rápido
Conservador	Agressivo



Exemplo

29/07/2014 19h15 - Atualizado em 29/07/2014 19h15

Número de usuários do Twitter sobe 24%, para 271 milhões

Por um lado, microblog registrou receita 124% maior, de US\$ 312 milhões. Por outro, o prejuízo no trimestre cresceu 242%, para US\$ 144,6 milhões.

Link da Matéria: <http://bit.ly/seminariobd>

Fonte: G1



NoSQL

- O que é?
- Quando Surgiu?
- Qual seu real Objetivo?



NoSQL

○ O que é?

- É a denominação genérica para banco de dados Não Somente relacionais (**Not Only**)

○ Quando surgiu?

- Apresentado por Carlo Strozzi, o termo foi usado em 1998 pela primeira vez para um BD que omitiu o uso do SQL e depois seria usado novamente em 2009, em duas conferências de defensores de BDs NoSQL

○ Qual o seu objetivo?

- Atender a necessidade de empresas que possuem um tráfego de dados muito grande e precisam de rápido desempenho no processamento desses dados



SGBDs NoSQL x SGBDs Relacionais

	NoSQL	SQL
Model	Non-relational Stores data in JSON documents, key/value pairs, wide column stores, or graphs	Relational Stores data in a table
Data	Offers flexibility as not every record needs to store the same properties	Great for solutions where every record has the same properties
	New properties can be added on the fly	Adding a new property may require altering schemas or backfilling data
	Relationships are often captured by denormalizing data and presenting it in a single record	Relationships are often captured in a using joins to resolve references across tables
	Good for semi-structured data	Good for structured data
Schema	Dynamic or flexible schemas Database is schema-agnostic and the schema is dictated by the application. This allows for agility and highly iterative development	Strict schema Schema must be maintained and kept in sync between application and database
Transactions	ACID transaction support varies per solution	Supports ACID transactions
Consistency	Consistency varies per solution, some solutions have tunable consistency	Strong consistency supported
Scale	Scales well horizontally	Scales well vertically



Principais modelos de dados em NoSQL

- Banco de dados orientado a Colunas
- Banco de dados orientado a Grafos
- Banco de dados orientado a Documentos
- Banco de dados Chave-Valor



Principais modelos de dados em NoSQL

- Banco de dados orientado a Colunas
- Banco de dados orientado a Grafos



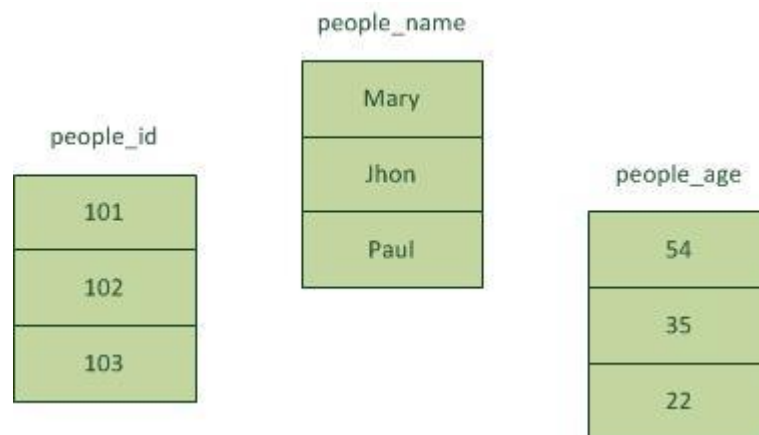
Banco de Dados Orientados a Colunas

- Criada para Processar grande quantidade de dados
- É contrário ao modelo relacional
- Não possuem Relacionamentos
- Não da suporte ao uso de chaves estrangeiras



Banco de Dados Orientado a Colunas

people_id	people_name	people_age
101	Mary	54
102	Jhon	35
103	Paul	22





Exemplos

- Big Table
- Cassandra
- Hbase



Vantagens do Modelo de Colunas

- Altamente Escalável
- Ordenação e Manutenção de Dados
- Maior Granularidade
- Indexação



Desvantagens do Modelo de Colunas

- Não é tão eficiente quando se trata de pouca quantidade de dados
- Recuperação complexa de registros



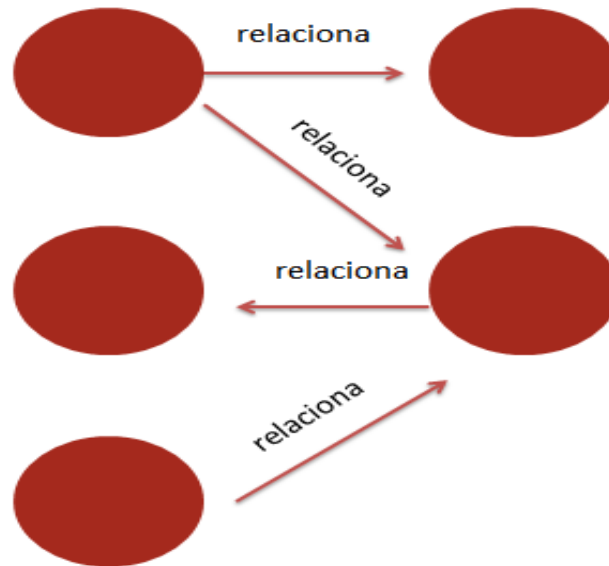
Banco de dados orientado a Grafos

- O que é?
 - Sistema que armazena dados em formas de grafos
 - Possui o foco direcionado apenas em relacionamentos



Banco de dados orientado a Grafos

Em um banco de dados de grafos, relacionamentos são mais naturais. Temos as entidades chamadas de vértices (ou node) que são ligadas entre elas pelas arestas (ou relationships) cada um podendo guardar dados entre os relacionamentos e cada relacionamento pode ter uma direção.





Itens de um banco de dados de Grafos

- Vértices ou Nó

- Unidade de Dados que representa uma entidade

- Arestas

- São os relacionamentos entre as Vértices, podendo ter um Sentido

- Propriedades

- Atributos das vértices e das Arestas



Banco de Dados Orientado a Grafos

- Exemplos de banco de dados de grafos:
 - Neo4J (O mais conhecido)
 - ArangoDb
 - HyperGraphDb
 - Oracle Spatial



Os SGBDs que as empresas usam:

- **Amazon Dynamo:** Amazon.
- **Apache CouchDB:** Apache.
- **SimpleDB:** Amazon.
- **Voldemort:** LinkedIn.
- **Cassandra:** Facebook, Twitter, e a Digg.
- **Google Bigtable:** Google, e a Caelum.
- **Mixi:** Facebook do Japão.
- **MongoDB:** New York Times.



Conclusões

- Os SGBDs NoSQL foram desenvolvidos para suprir a necessidade de processar grande quantidade de dados.
- No Modelo de Colunas, armazenamos várias colunas de acordo com a necessidade de Inserirmos novos dados
- Também podemos alterar somente uma coluna ao invés de termos que alterar uma tupla inteira como no modelo relacional, ganhando tempo e economizando Hardware



Conclusões

- O modelo de grafos é comparado com um multigrafo direcionado e cada nó pode ser conectado por mais de uma aresta
- É composto por três componentes básicos: Vértices, Arestas e as propriedades, que são os atributos das Vértices e das Arestas



Referências

<http://www.100loop.com/nosql-quando-e-qual-usar/>

<https://pt.stackoverflow.com/questions/122452/como-escolher-entre-nosql-e-sql>

<http://www.devmedia.com.br/o-que-e-nosql-java-magazine-86/18777>

<http://www.itexto.net/devkico/?p=1199>

<https://pt.wikipedia.org/wiki/NoSQL>

<https://isaiasbarroso.wordpress.com/2012/06/20/banco-de-dados-orientado-a-colunas/>

<https://imasters.com.br/banco-de-dados/graphdb-series-o-que-e-um-banco-de-dados-de-grafos/?trace=1519021197&source=single>

<https://medium.com/labcodes/graph-databases-discutindo-o-relacionamento-dos-seus-dados-com-python-79688b557eec>

<http://blog.caelum.com.br/trabalhando-com-relacionamentos-bancos-de-dados-baseados-em-grafos-e-o-neo4j/>

<http://g1.globo.com/tecnologia/noticia/2014/07/numero-de-usuarios-do-twitter-sobe-24-para-271-milhoes.html>



Perguntas

