



Universidade de Brasília

**Instituto de Ciências Exatas
Departamento de Ciência da Computação**

Armazenamento de Dados Abertos com NoSQL

Jorge Luiz Andrade

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientador

Prof.^a Dr.^a Maristela Terto de Holanda

Brasília
2016

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Coordenador: Prof. Dr. Flávio de Barros Vidal

Banca examinadora composta por:

Prof.^a Dr.^a Maristela Terto de Holanda (Orientador) — CIC/UnB
Prof. Dr. Professor I — CIC/UnB
Prof. Dr. Professor II — CIC/UnB

CIP — Catalogação Internacional na Publicação

Andrade, Jorge Luiz.

Armazenamento de Dados Abertos com NoSQL / Jorge Luiz Andrade.
Brasília : UnB, 2016.

21 p. : il. ; 29,5 cm.

Monografia (Graduação) — Universidade de Brasília, Brasília, 2016.

1. bancos de dados, 2. nosql, 3. dados abertos

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Armazenamento de Dados Abertos com NoSQL

Jorge Luiz Andrade

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Prof.^a Dr.^a Maristela Terto de Holanda (Orientador)
CIC/UnB

Prof. Dr. Professor I Prof. Dr. Professor II
CIC/UnB CIC/UnB

Prof. Dr. Flávio de Barros Vidal
Coordenador do Bacharelado em Ciência da Computação

Brasília, 15 de dezembro de 2016

Dedicatória

Dedico a....

Agradecimentos

Agradeço a....

Abstract

A ciência...

Palavras-chave: bancos de dados, nosql, dados abertos

Abstract

The science...

Keywords: databases, nosql, open data

Sumário

| | | |
|----------|-------------------------|----------|
| 1 | Introdução | 1 |
| 1.1 | NoSQL | 1 |
| 1.2 | Dados Abertos | 2 |
| | Referências | 3 |

Lista de Figuras

Lista de Tabelas

| | |
|---------------------------------------|---|
| 1.1 Modelos de Bancos NoSQL | 2 |
|---------------------------------------|---|

Capítulo 1

Introdução

Podemos definir um banco de dados relacional como um conjunto de dados que se relacionam entre si, armazenados de uma forma persistente, ou seja, que possam ser recuperados quando necessários. Devido às suas características de persistência, concorrência, integração e padronização, tem sido o modelo padrão de armazenamento, principalmente em ambientes empresariais, a pelo menos vinte anos na computação[5]. Uma questão importante em bancos de dados computacionais atuais é a não necessidade que o usuário tem de conhecer como esses dados são armazenados. Isso é possível graças aos chamados Sistemas Gerenciadores de Bancos de Dados(SGBDs)[4].

1.1 NoSQL

Nos últimos anos, devido a novas necessidades, tem surgido um movimento em direção a um novo paradigma denominado NoSQL. O termo foi utilizado pela primeira vez em 1998 para denominar um banco de dados que omitia o uso de SQL. A definição atual, porém, tem suas bases em uma reunião realizada em 2009 em São Francisco, Estados Unidos. Organizada por Johan Oskarsdon, criador do Last.fm, nela foram discutidas formas mais eficientes e baratas de organização dos dados, como as já sugeridas pelo Google Bigtable e Amazon's Dynamo em publicações anteriores[6].

O termo não tem uma definição precisa e universalmente aceita, mas é geral descrito como "Not Only SQL". Bancos NoSQL em geral são caracterizados, mas não definidos, como sendo não relacionais, sem esquema bem definido e distribuídos, favorecendo a execução em clusters, apesar de existirem exceções, como os bancos de dados de grafos, que são executados geralmente em um único servidor.

Bancos de dados NoSQL possuem padrões de modelos de dados, que compartilham certas características em comum e servem a determinadas aplicações específicas, podendo alguns bancos serem classificados em mais de uma categoria. A tabela 1.1 lista os quatro modelos atuais e alguns bancos de dados que se enquadram em cada um deles.

Tabela 1.1: Modelos de Bancos NoSQL

| Modelo de Dados | Exemplo de bancos de dados |
|------------------------|---|
| Chave-valor | Project Voldemort Riak BerkeleyDB |
| Documentos | CouchDB MongoDB OrientDB |
| Famílias de colunas | Cassandra HBase |
| Grafos | Neo4j, OrientDB, Infinite Graph OrientDB Infinite Graph |

1.2 Dados Abertos

A Open Definition define um dado como aberto "se qualquer pessoa esta livre para acessá-lo, utilizá-lo, modificá-lo, e compartilhá-lo — restrito, no máximo, a medidas que preservam a proveniência e abertura." [1].

O poder público brasileiro vem nos últimos anos realizando ações que promovem a abertura de dados governamentais. Essas ações visam benefícios como melhoria da gestão pública, transparência, controle a participação social, geração de emprego e renda e estímulo à inovação tecnológica. [3]. Para atingir esse fim, no ano de 2012 foi definido, em instrução normativa, a implantação da INDA, Infraestrutura Nacional de Dados Abertos, "um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos" [2].

Referências

- [1] Definição de conhecimento aberto. <http://opendefinition.org/od/2.0/pt-br/>. Acessado em 08 de abril de 2016. 2
- [2] Inda - infraestrutura nacional de dados abertos. <http://www.governoeletronico.gov.br/acoes-e-projetos/Dados-Abertos/inda-infraestrutura-nacional-de-dados-abertos>. Acessado em 10 de abril de 2016. 2
- [3] Tribunal de Contas da União. 5 motivos para abertura de dados na administração pública. <http://portal3.tcu.gov.br/portal/pls/portal/docs/2689107.PDF/>. 2
- [4] Jan L. Harrington. *Relational database design and implementation : clearly explained*. Morgan Kaufmann, 2009. 1
- [5] Pramod J. Sadalage and Martin Fowler. *NoSQL Essencial*. Novatec, 2013. 1
- [6] Christof Strauch. *Nosql databases*. Stuttgart Media University, 2011. 1