

What is NoSQL?

Mestrado Integrado em Engenharia Informática

Unidade Curricular Completar

What is NoSQL?

António Abelha

Departamento de Informática

Escola de Engenharia



What is NoSQL?

NoSQL é um termo que se refere a um tipo específico de modelo de bases de dados ou sistema de gestão de bases de dados (DBMS).

O termo **NoSQL** é muito amplo, que não se refere a um modelo específico de bases de dados. Refere-se a toda uma panóplia de diferentes modelos que têm como principal característica não se encaixarem no modelo relacional.

Embora as bases de dados NoSQL existam desde os anos 60 do século passado, foi apenas no início deste século que a abordagem NoSQL começou a ganhar força, e toda uma nova geração de sistemas NoSQL começou a chegar ao mercado.



What is NoSQL?

Um problema chamado **Big Data**

Uma das principais razões pelas quais a abordagem NoSQL começou a ser adoptada ocorreu devido ao *big data*.

Empresas como Google e Amazon começaram a ter grandes dificuldades a lidar com volumes de dados crescentes devido à sua crescente popularidade.

Isso levou ao seguinte:

What is NoSQL?

A Google desenvolveu o *Bigtable*

Um sistema de armazenamento distribuído para gerir dados estruturados, projetada para funcionar em segurança para petabytes de dados e milhares de máquinas.

O *Bigtable* é usado em dezenas produtos e projetos, incluindo Pesquisa, Analytics, Mapas e Gmail.

O objetivo do *Bigtable* é fornecer ampla aplicabilidade, escalabilidade, alto desempenho e alta disponibilidade.

(ver documento sobre Bigtable no *elearning*).



What is NoSQL?

A Amazon desenvolveu e implementou o *Dynamo*

Um sistema de bases de dados do tipo Key-Values (valor-chave) com alta disponibilidade que alguns dos principais serviços da Amazon usam para proporcionar ao utilizadores uma experiência designada por "sempre ativa". Para poder atingir um tão alto grau de disponibilidade, o *Dynamo* sacrifica a consistência em cenários adversos e de falha.

(ver documento sobre *Dynamo* no elearning)



What is NoSQL?

Mas a Google e a Amazon não são as únicas empresas que lidam com o problema do *big data*.

Muitas outras desenvolveram soluções inovadoras por razões semelhantes.

A inovação neste campo continua ainda hoje.

Temos tipo nos últimos anos uma explosão de novas tecnologias desenvolvidas fora da estrutura relacional das DBMS

“NoSQL”.



What is NoSQL?

Definição de *NoSQL*

????????? Muitas se discute sobre este assunto ????????????????

"**Sem SQL**" (estes sistemas não utilizam **SQL** - usam uma linguagem de consulta alternativa).

"**Não apenas SQL**" (sistemas que usam **SQL** em conjunto com outras tecnologias / outras linguagens de consulta).

Muitos argumentam que a única coisa que todas as bases de dados NoSQL têm em comum é que não seguem o modelo relacional.

"**NoREL**" seria um nome mais adequado.



What is NoSQL?

Características de uma base de dados **NoSQL**

O termo NoSQL refere-se a um grupo específico de **DBMSs** que partilham certas características:

- Não Relacionais
- Código aberto
- Sem esquema
- Escaláveis horizontalmente
- Não seguem os princípios do **ACID**
- Sem uma linguagem de consulta *padrão*.

Nem todas as bases de dados **NoSQL** possuem estas características.

No entanto, a maioria dessas características é inerentemente inexistente nas bases de dados relacionais.

What is NoSQL?

Modelo Não Relacional

Modelos de bases de dados não baseados no modelo relacional proposto por E.F. Codd em 1970.

DBMSs não relacionais porque na maior parte das situações o modelo relacional não é o mais adequado para os requisitos.

What is NoSQL?

Modelo Não Relacional

A situação mais comum ocorre nos casos em que os dados sejam não estruturados ou semiestruturados.

O volume de dados (big data) pode exigir uma nova maneira de abordar o problema do armazenamento e da recuperação desses dados.

Muitas vezes as organizações podem precisar de escalar os seus sistemas para centenas ou milhares de computadores.

Podemos ter muitas outras razões.

What is NoSQL?

Código aberto

A maioria dos ***DBMSs*** *NoSQL* é disponibilizado em código aberto.

Embora também existam muitos DBMSs modelo relacionais em código aberto.

Quase todo o movimento NoSQL tende a orientar-se para projetos de código aberto, com muitas organizações contribuindo para os esforços de desenvolvimento de uma única solução.

Código aberto não é necessariamente um "requisito NoSQL", mas é uma "observação NoSQL".



What is NoSQL?

Sem esquema

A maioria dos sistemas de bases de dados NoSQL não possui um esquema fixo.

Nas DBMS modelo relacional temos sempre de considerar o esforço de modelação antes da introdução de quaisquer dados.

Numa base de dados **NoSQL** não necessitamos desse requisito.

Neste contexto os modelos noSQL tendem a ser mais flexíveis com os dados que pode aceitar, tornando todo o processo de desenvolvimento muito ágil.

What is NoSQL?

Escaláveis horizontalmente

A maioria das bases de dados NoSQL tem um comportamento excelente em ambientes cluster.

Aqui que os dados são particionados em vários modos para que cada todo possa executar uma tarefa específica e muitas vezes independentemente.

Cada todo pode executar sua tarefa sem a necessidade de partilhar memória ou espaço em disco com outros processos.

Processo conhecido por *shared nothing architecture* (SN).

What is NoSQL?

Escaláveis horizontalmente

As bases de dados relacionais também podem ser configurados para funcionar em ambiente *cluster*, a abordagem *RDBMS* tende a dificultar a sua configuração, quando comparada com as bases de dados *NoSQL*.

O desempenho também pode ficar comprometido quando escalamos uma base de dados relacional dessa maneira.

As bases de dados relacionais estão vocacionadas mais para a “expansão”, para adicionar mais recursos a uma única máquina de modo a obtermos uma máquina com melhor desempenho.

What is NoSQL?

Não seguem os princípios do ACID

A grande maioria das bases de dados NoSQL

Não é rigorosa no cumprimento as restrições de ACID (Atomicidade, Consistência, Isolamento, Durabilidade), em diferentes graus.

What is NoSQL?

ACID (**A**tomicidade, **C**onsistência, **I**solamento e **D**urabilidade)
(*Atomicity, Consistency, Isolation, Durability*)

é o conjunto das propriedades de uma transação numa bases de dados.



What is NoSQL?

ACID (**A**tomicidade, **C**onsistência, **I**solamento e **D**urabilidade)
(*Atomicity, Consistency, Isolation, Durability*)

Atomicidade:

Uma transação deve ter todas as suas operações executadas em caso de sucesso, em caso de falha, nenhum resultado de alguma operação refletido sobre a base de dados.



What is NoSQL?

ACID (**A**tomicidade, **C**onsistência, **I**solamento e **D**urabilidade)
(*Atomicity, Consistency, Isolation, Durability*)

Consistência:

A execução de uma transação deve levar a base de dados de um estado consistente a um outro estado consistente,

Uma transação deve respeitar as regras de integridade dos dados (como unicidade de chaves, restrições de integridade lógica, etc.).



What is NoSQL?

ACID (**A**tomicidade, **C**onsistência, **I**solamento e **D**urabilidade)
(*Atomicity, Consistency, Isolation, Durability*)

Isolamento:

O *isolamento* é um conjunto de técnicas que evitam que transações paralelas interfiram umas nas outras, como se todas essas transações fossem executadas sequencialmente (uma após a outra). Operações exteriores a uma dada transação jamais verão esta transação em estados intermediários.

Fornecer isolamento é o objetivo principal do controle da concorrência.

What is NoSQL?

ACID (**A**tomicidade, **C**onsistência, **I**solamento e **D**urabilidade)
(*Atomicity, Consistency, Isolation, Durability*)

Durabilidade:

O resultado da execução de uma transação em caso de sucesso (*commit*) devem persistir na base de dados mesmo em casos de falta de energia, bloqueamentos ou erros.

Devem garantir que os dados estarão disponíveis em definitivo.



What is NoSQL?

Não seguem os princípios do ACID

A porque a maioria das soluções NoSQL foi desenvolvida com o objetivo de fornecer alta disponibilidade e escalabilidade em ambiente *cluster*.

Ao fornecer alta disponibilidade em ambiente *cluster*, é necessário sacrificar a consistência ou a durabilidade (ou encontrar um equilíbrio entre os três).

Ao “facilitar” na consistência da informação, é possível obter uma maior disponibilidade.

What is NoSQL?

Não seguem os princípios do ACID

A abordagem NoSQL reconhece que existem alguns casos em que manter dados inconsistentes não é o “*fim do mundo*”.

Desde que tal seja assumido e essa inconsistência possa ser tratada de alguma forma, não é crítico apresentar dados que não sejam 100% consistentes.

No entanto, alguns sistemas NoSQL, como MarkLogic, Neo4j e OrientDB, suportam e respeitam as transações ACID.

What is NoSQL?

Sem uma linguagem de consulta *padrão*.

Ainda não temos uma linguagem de consulta padrão que seja suportada por todas as bases de dados NoSQL.

Embora já tenha havido tentativas de introduzir uma linguagem de consulta padrão para todas as bases de dados NoSQL, nenhuma foi adoptada.

Alguns DBMSs *NoSQL* têm sua própria linguagem de consulta, enquanto outros suportam várias linguagens, como **JSON**, **XQuery**, **SPARQL**, etc.

What is NoSQL?

As primeiras bases de dados NoSQL

Muitas bases de dados "não SQL" foram propostas ao longo dos anos. Antes mesmo do SQL ser proposto pela primeira vez em 1974. O primeiro RDBMS apenas foi lançado comercialmente em 1979.

Em seguida os três dos primeiros sistemas de gestão de bases de dados (NoSQL) apresentados:

What is NoSQL?

As primeiras bases de dados NoSQL

O ***MultiValue*** é uma base de dados NoSQL apresentada pela primeira vez em 1965 como parte do sistema operativo Pick.

O ***MUMPS*** (*Massachusetts General Hospital Utility Multi-Programming System*) desde 1966. *MUMPS* não tem esquema, e usa um motor de base de dados do tipo *key-values*, é um exemplo clássico de uma base de dados *NoSQL*.

O ***IBM IMS*** é a junção entre uma base de dados hierárquica e um sistema de gestão de informações desenvolvido em 1996.

What is NoSQL?

As primeiras bases de dados NoSQL

Na altura ninguém chamou a estas bases de dados “**NoSQL**”.
No caso das duas primeiras o **SQL** ainda não tinha sequer sido inventado.

Só depois de 2009 é que estes **DBMSs** mais antigos foram associados ao movimento **NoSQL**.

What is NoSQL?

Tipos de bases de dados *NoSQL*

A bases de dados NoSQL são actualmente classificadas em quatro tipos principais.

Algumas delas de dados são uma mistura entre diferentes tipos diferentes.
Em geral todas se encaixam nestas categorias.

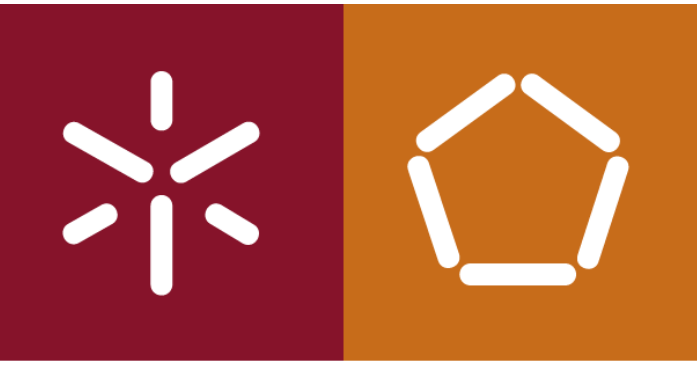


What is NoSQL?

O que é uma Key-Value Database?

*A key-value database (também conhecida por *key-value store* e *key-value store database*)*

É um tipo de base de dados *NoSQL* que usa um método simples chave / valor para armazenar dados.



What is NoSQL?

A parte do **key-value** refere-se ao fato de a base de dados armazenar dados como uma coleção de pares de chave / valor.

É um método muito simples de armazenamento de dados, conhecido por estar muito bem dimensionado, por rentabilizar os espaço ocupado, evitando colunas a nulo.

O par de valores-chave é um conceito bem estabelecido em muitas linguagens de programação.

Muitas linguagens de programação quando se referem a um valor-chave identificam como uma matriz dispersa ou uma estrutura de dados.

Um valor-chave é também muitas vezes identificado por dicionário ou por *hash table*.

What is NoSQL?

Que tipos de dados podem ser armazenados numa base de dados valores-chave?

A chave

A chave deve ser exclusiva. É o identificador exclusivo que permite aceder ao valor associado a essa chave.

A chave pode ser qualquer coisa. Mas muitas vezes depende da DBMS.
Um DBMS pode impor limitações, enquanto outro pode não impor nenhuma.

What is NoSQL?

Que tipos de dados podem ser armazenados numa base de dados valores-chave?

No **Redis**, por exemplo, o tamanho máximo permitido para a chave é de 512 MB. Podemos usar uma sequência binária como chave, Uma sequência de texto associado ao conteúdo de uma de imagem. Até uma sequência vazia é uma chave válida.

What is NoSQL?

Que tipos de dados podem ser armazenados numa base de dados valores-chave?

No entanto, por razões de desempenho, devemos evitar uma chave muito longa. Muito curta pode causar problemas de legibilidade. A chave deve seguir uma convenção acordada para manter os dados consistentes.

What is NoSQL?

O valor

O valor armazenado pode ser qualquer coisa, como texto (longo ou curto), um número, código HTML, JSON, código de programação PHP, uma imagem etc.

O valor também pode ser uma lista ou até um outro par *Key-values* encapsulado num objeto.

What is NoSQL?

O valor

Alguns DBMSs permitem especificar tipos de dados para o valor. Podemos especificar que o valor deve ser um número inteiro. Outros DBMSs não incorporam essa funcionalidade, o valor pode ser de qualquer tipo.

What is NoSQL?

O valor

O **Redis DBMS** permite especificar os seguintes tipos de dados:

- **Cadeias binárias** seguras.
- **Listas**: coleções de elementos de sequência classificados de acordo com a ordem de inserção.
- **Conjuntos**: coleções de elementos de uma sequência, exclusivos e não classificados.
- **Conjuntos classificados**, semelhantes aos Conjuntos, mas onde cada elemento da cadeia está associado a um valor numérico flutuante, score. Para operações do tipo, selecionar os 10 melhores ou os 10 inferiores, etc.
- **Hashes**, mapas compostos por campos associados a valores. O campo e o valor são cadeias de caracteres.
- **Matrizes de bits** (ou simplesmente bitmaps).
- **HyperLogLogs**: é uma estrutura de dados probabilísticos usados para estimar a cardinalidade de um conjunto.

What is NoSQL?

Exemplos de utilização de bases de dados Key-Values?

As bases de dados de valores-chave podem ser utilizadas em muitos cenários.

- Aplicações *web*
 - Perfis de utilizador
 - Informações da sessão
 - Artigos / *blogs* de comentários
 - Gestão de *Emails*
 - Mensagens de estado

What is NoSQL?

- **Comércio electrónico**

- Conteúdo dos carrinhos de compras
- Categorias de Produtos
- Detalhes do produto
- Consultas de produtos

- **Redes / Manutenção de Dados**

- Empresas de telecomunicações
- Tabelas de encaminhamento de protocolo Internet (IP)
- Deduplicação de dados (técnica para eliminar cópias duplicadas em dados repetidos)

What is NoSQL?

As bases de dados de valores-chave podem armazenar páginas Web, Em que o URL como chave e a página como valor.

Algumas das DBMSs que utilizam uma abordagem de valor-chave.

- Redis
- Base de Dados Oracle NoSQL
- Voldemorte
- Aerospike
- Oracle Berkeley DB

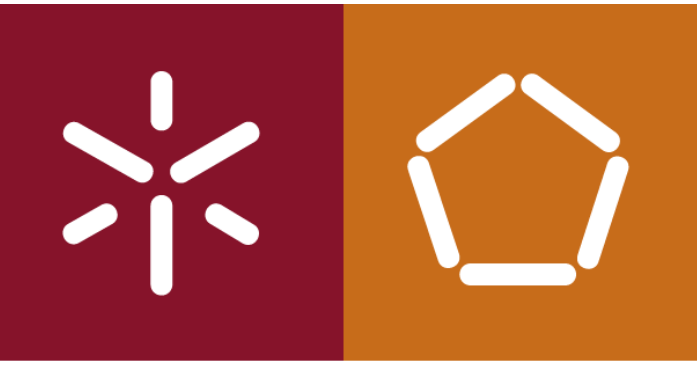
What is NoSQL?

Top 8 NoSQL Key value Database

[https://www.predictiveanalyticstoday.com/top-sql-key-value-store-databases/
#comments](https://www.predictiveanalyticstoday.com/top-sql-key-value-store-databases/#comments)

What is NoSQL?

O que é uma Document Store Database?



What is NoSQL?

Uma base de dados de armazenamento de documentos

- Bases de dados orientadas a documentos,
- Bases de dados agregadas
- Armazenamento de documentos
- Bases de dados documentais)

É uma base de dados que usam modelos orientados a documentos para armazenar dados.

What is NoSQL?

As bases de dados documentais armazenam cada registo e todos os dados associados num único documento.

Cada documento é composto por dados semiestruturados.

Dados esses que podem ser consultados usando várias ferramentas de consulta e análise do DBMS.

What is NoSQL?

Exemplos de documentos

Em seguida dois exemplos de documentos que podem ser armazenados numa base de dados documental.

Os mesmos dados apresentados em *idiomas diferentes*.

What is NoSQL?

XML.

```
<artist>
  <artistname>Iron Maiden</artistname>
  <albums>
    <album>
      <albumname>The Book of Souls</albumname>
      <datereleased>2015</datereleased>
      <genre>Hard Rock</genre>
    </album>
    <album>
      <albumname>Killers</albumname>
      <datereleased>1981</datereleased>
      <genre>Hard Rock</genre>
    </album>
    <album>
      <albumname>Powerslave</albumname>
      <datereleased>1984</datereleased>
      <genre>Hard Rock</genre>
    </album>
    <album>
      <albumname>Somewhere in Time</albumname>
      <datereleased>1986</datereleased>
      <genre>Hard Rock</genre>
    </album>
  </albums>
</artist>
```

What is NoSQL?

JSON.

```
{
  '_id' : 1,
  'artistName' : { 'Iron Maiden' },
  'albums' : [
    {
      'albumname' : 'The Book of Souls',
      'datereleased' : 2015,
      'genre' : 'Hard Rock'
    }, {
      'albumname' : 'Killers',
      'datereleased' : 1981,
      'genre' : 'Hard Rock',
    }, {
      'albumname' : 'Powerslave',
      'datereleased' : 1984,
      'genre' : 'Hard Rock'
    }, {
      'albumname' : 'Somewhere in Time',
      'datereleased' : 1986,
      'genre' : 'Hard Rock'
    }
  ]
}
```

What is NoSQL?

Exemplos de documentos

Adicionamos o campo *_id* no segundo exemplo (**JSON**).

Nem sempre isso é obrigatório.

Alguns **DBMSs** inserem automaticamente um **ID** exclusivo.



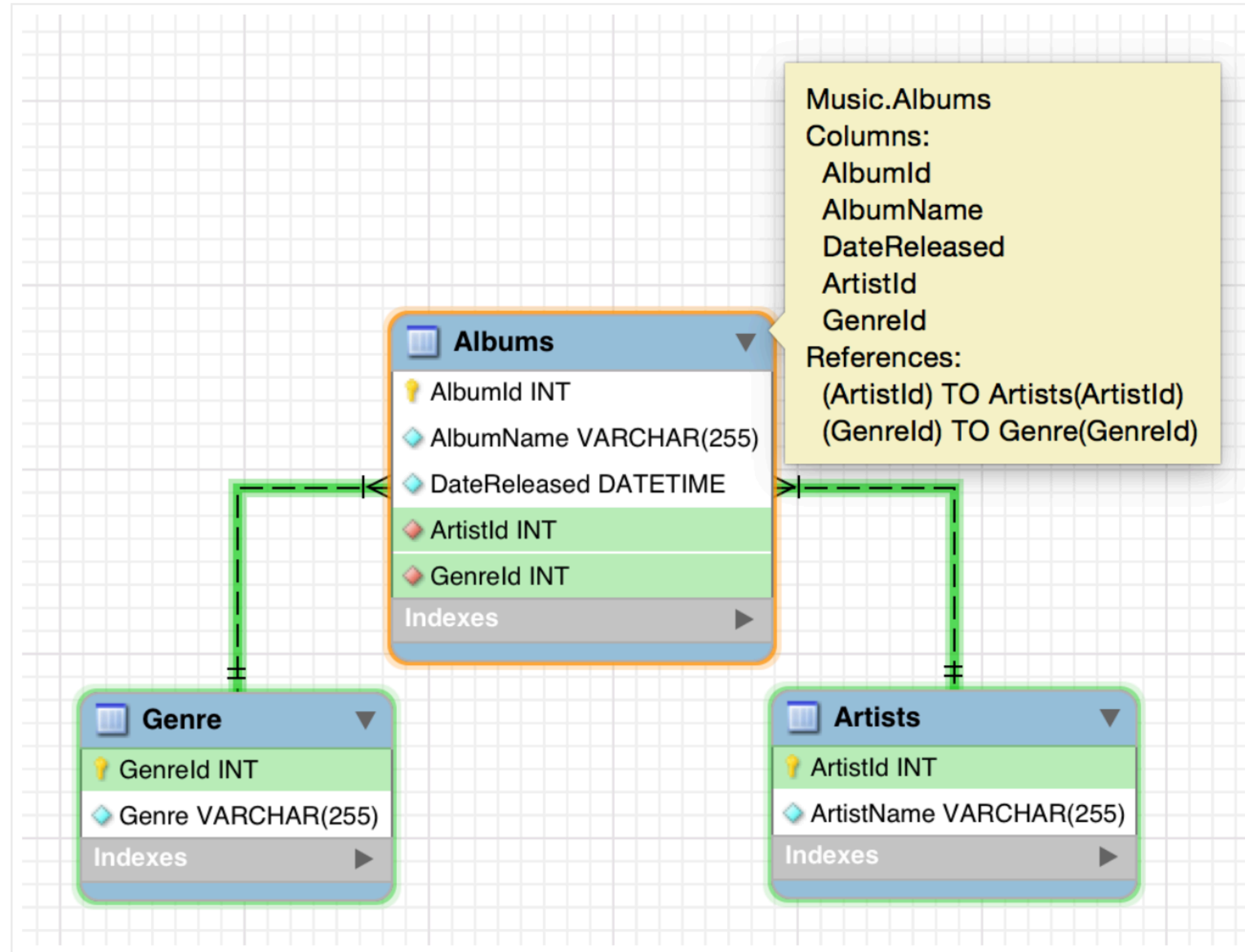
What is NoSQL?

Documento Store vs Bases de dados relacionais

Se inserirmos os dados apresentados numa base de dados relacional, se normalizarmos a informação vamos ter de utilizar três tabelas diferentes.

Relacionadas por meio dos atributos chave primária e chave estrangeira.

What is NoSQL?



What is NoSQL?

Artists

ArtistId	ArtistName
1	Iron Maiden
2	Devin Townsend
3	The Wiggles
4	...

Albums

AlbumId	AlbumName	DateReleased	ArtistId	GenreId
1	The Book of Souls	2015	1	3
2	Killers	1981	1	3
3	Powerslave	1984	1	3
4	Somewhere in Time	1986	1	3
5	Ziltoid the Omniscient	2007	2	3
6

Genre

GenreId	Genre
1	Country
2	Blues
3	Hard Rock
4	...

What is NoSQL?

Tabelas

A bases de dados relacionais armazenam dados em tabelas, Cada tabela contem colunas e cada linha representa um registo. As informações sobre uma entidade podem ser espalhadas por várias tabelas. Os dados de tabelas diferentes podem estar associados quando estabelecemos um relacionamento entre as tabelas.

As bases de dados documentais, não usam tabelas. Os dados de uma determinada entidade são armazenados num único documento e todos os dados associados são armazenados dentro desse documento.

What is NoSQL?

Esquemas

Nas bases de dados relacionais, criamos um esquema lógico antes de inserirmos qualquer dado.

Nas bases de dados documentais (característica comum com grande parte das bases de dados **NoSQL**), não temos esse requisito.

Podemos simplesmente carregar todos os dados sem usar qualquer esquema predefinido.

What is NoSQL?

Escalabilidade

As bases de dados de documentais podem ser dimensionadas horizontalmente.

Os dados podem ser armazenados em milhares de computadores e o sistema terá sempre um bom desempenho.

sharding.

What is NoSQL?

Escalabilidade

A bases de dados relacionais não são adequadas este tipo de utilização.

As bases de dados relacionais são mais adequados um dimensionamento vertical. (i.e., adicionar mais memória, espaço de disco etc.).

Temos sempre um limite para os recursos de uma máquina,
A escalabilidade horizontal pode ser a única opção.

What is NoSQL?

Relacionamentos

Os dados armazenados nas bases de dados documentais não possuem chaves estrangeiras

As chaves estrangeiras são usadas nas bases de dados relacionais para impor relacionamentos entre tabelas.

Nível aplicativo apenas

What is NoSQL?

Relacionamentos

A ideia por trás do modelo de documental é a que todos os dados associados a um registo são armazenados no mesmo documento.

A necessidade de estabelecer um relacionamento no modelo de documental não tem o mesmo impacto que numa base de dados relacional.

What is NoSQL?

NoSQL

As bases de dados relacionais usam SQL como a linguagem de interrogação.

As bases de dados documentais tendem a usar outras linguagens de *query* (embora alguns ofereçam suporte para SQL).

A maioria disponibilizam linguagens como XQuery, XSLT, SPARQL, Java, JavaScript, Python, etc.



What is NoSQL?

Document Store vs Key-Value Databases

As bases de dados de documentais são semelhantes às bases de dados valor-chave

(uma chave -> um valor)

Os dados são armazenados como um valor.

A chave associada é um identificador exclusivo para esse valor.



What is NoSQL?

Document Store vs Key-Value Databases

Numa base de dados documental, o valor contém dados estruturados ou semiestruturados.

Os valor **estruturado / semiestruturado** é um documento.

Os dados estruturados / semiestruturados que compõem um documento podem ser codificados em *XML*, *JSON*, *YAML*, *BSON* etc. Também podemos utilizar binários, como *PDFs*, documentos *MS Office*, etc.



What is NoSQL?

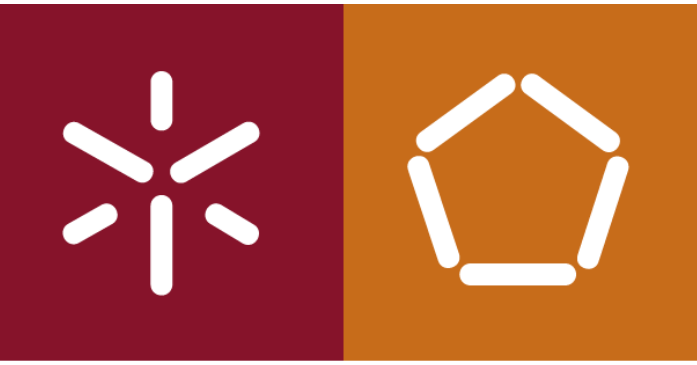
Vantagens dos Document Model sobre Key-Value Stores

A principal vantagem que as bases de dados documentais têm sobre as Key-Value Stores é podermos consular os dados em si.

Podemos consultar a estrutura do documento, bem como os elementos dessa estrutura.

Pode retornar apenas partes do documento.

Com uma base de dados Key-Value, acedemos a todo o valor, sendo ele ou não estruturado.



What is NoSQL?

Exemplos de Utilização dos Document Model Stores

Web Applications

- Content management systems
- Blogging platforms
- eCommerce applications
- Web analytics
- User preferences data

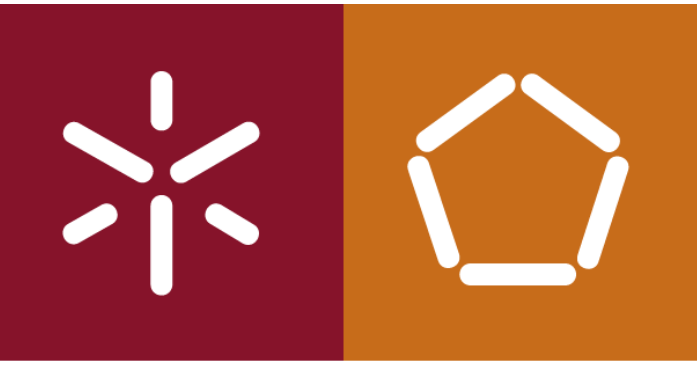


What is NoSQL?

Exemplos de Utilização dos Document Model Stores

User Generated Content

- Chat sessions
- Tweets
- Blog posts
- Ratings
- Comments



What is NoSQL?

Exemplos de Utilização dos Document Model Stores

Catalog Data

- User accounts
- Product catalogs
- Device registries for Internet of Things
- Bill of materials systems

What is NoSQL?

Exemplos de Utilização dos Document Model Stores

Gaming

- In-game stats
- Social media integration
- High-score leaderboards
- In-game chat messages
- Player guild memberships
- Challenges completed



What is NoSQL?

Exemplos de Utilização dos Document Model Stores

Networking/computing

- Sensor data from mobile devices
- Log files
- Realtime analytics
- Various other data from Internet of Things

What is NoSQL?

Exemplos de Document Model Stores

- [MongoDB](#)
- [DocumentDB](#)
- [CouchDB](#)
- [MarkLogic](#)
- [OrientDB](#)

What is NoSQL?

TOP 12 de Document Model Stores

<https://www.predictiveanalyticstoday.com/top-nosql-document-databases/>

What is NoSQL?

What is a Column Store Database?



What is NoSQL?

What is a Column Store Database?

Um base de dados de armazenamento de colunas é um tipo de base de dados que armazena dados utilizando um modelo de dados orientado a colunas.



What is NoSQL?

Uma base de dados para o armazenamento de colunas pode ser identificado por:

Base de dados de coluna

Base de dados de família de colunas

Base de dados orientada a colunas

Base de dados ampla para o armazenamento de colunas

Armazenamento de colunas alargadas

Base de dados colunar

Loja colunar



What is NoSQL?

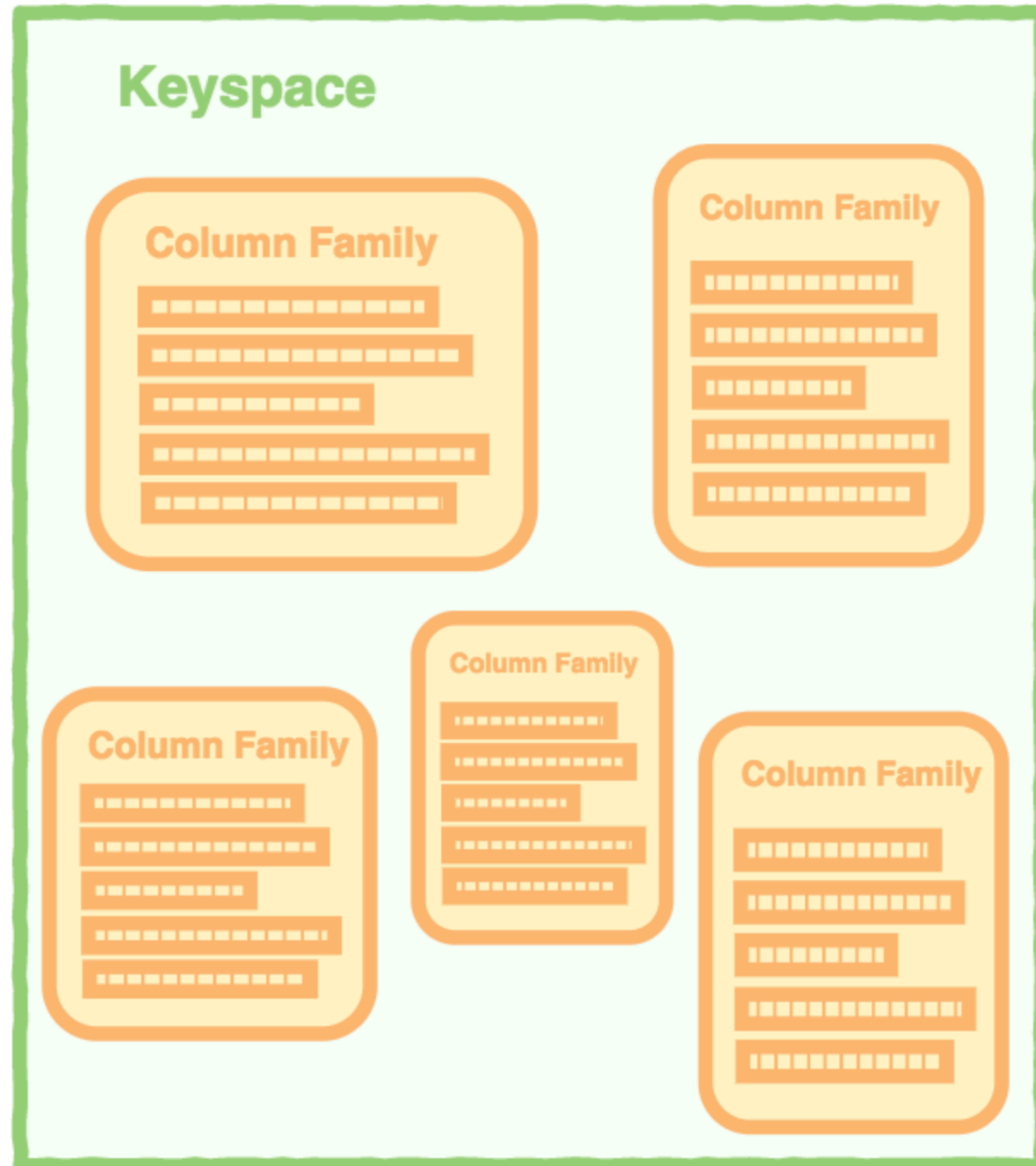
A estrutura de uma base de dados de armazenamento de colunas

As bases de dados de armazenamento de colunas usam um conceito chamado espaço de chave (*keyspace*).

Um espaço de chave é como um esquema no modelo relacional.

O espaço de chave contém todas as famílias de colunas (tabelas semelhantes às usadas no modelo relacional), que contêm linhas e colunas.

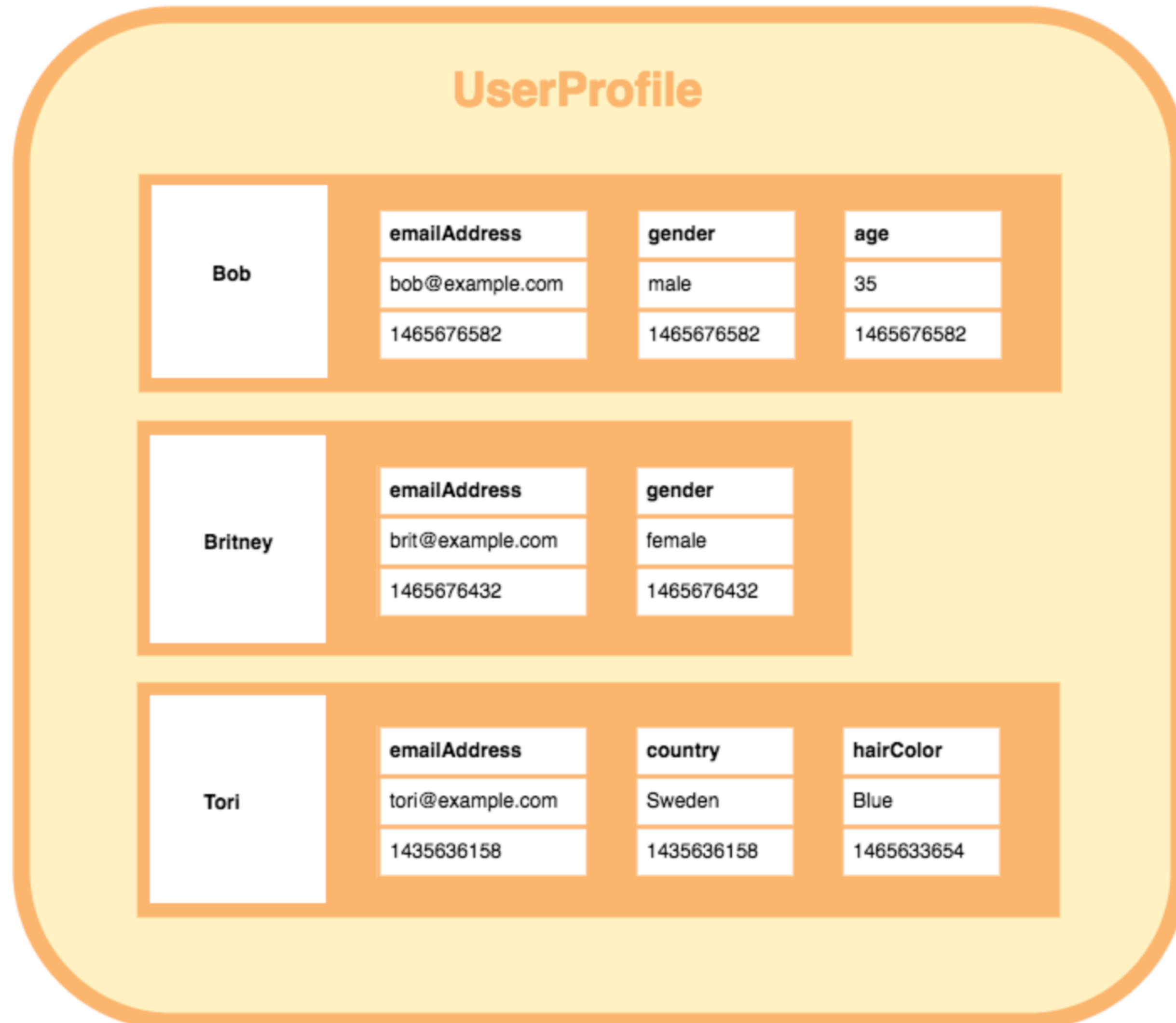
What is NoSQL?



Um ***Keyspace***
Com famílias de colunas.

What is NoSQL?

Uma Família



Esta família é composta por três linhas.
Cada linha contém seu próprio conjunto de colunas.



What is NoSQL?

Como podem ver

A família de colunas é composta por várias linhas.

Cada linha pode conter um número diferente de colunas.

As colunas não precisam corresponder às colunas de outras linhas (ou seja, podem ter diferentes nomes, tipos de dados etc.).

Cada coluna está contida na sua linha.

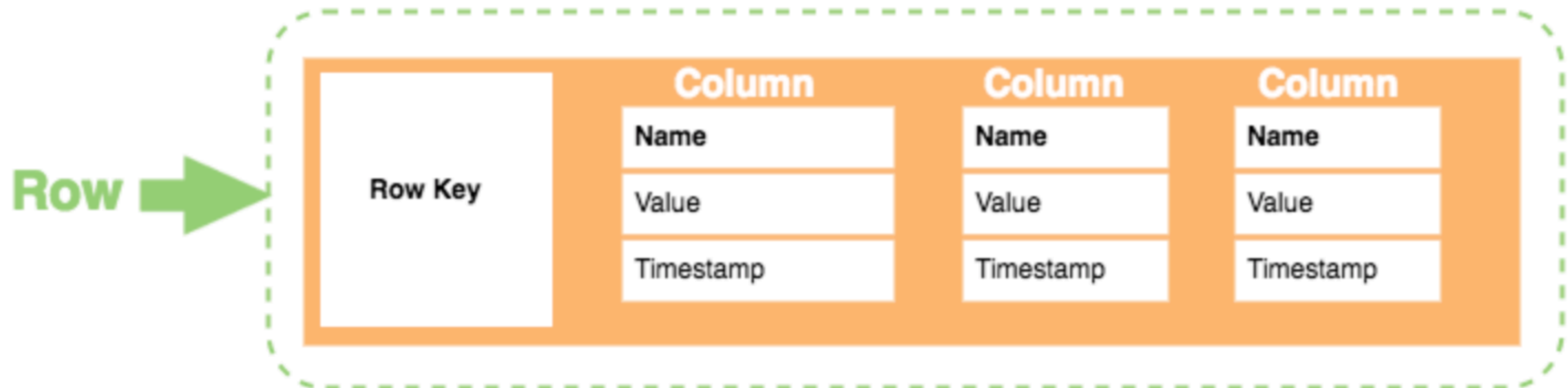
Não abrange todas as linhas, como acontece com uma base de dados relacional.

Cada coluna contém um par de nome / valor, juntamente com um carimbo de data / hora.

Nesste exemplo é utilizado o tempo Unix/Epoch para par data e hora.

What is NoSQL?

Elementos de uma linha





What is NoSQL?

Detalhe de cada elemento de uma linha:

Chave de linha: Cada linha possui uma chave exclusiva, que é um identificador exclusivo para essa linha.

Coluna: Cada coluna contém um nome, um valor e um carimbo de data / hora.

Nome: O nome do par nome / valor.

Valor: O valor do par nome / valor.

Registo de data e hora: Fornece a data e hora em que os dados foram inseridos. Isso pode ser usado para determinar a versão mais recente dos dados.

What is NoSQL?

Alguns DBMSs expandem o conceito de família de colunas para fornecer uma capacidade extra de armazenamento.

O ***Cassandra*** tem o conceito de colunas compostas, que permitem aninhar objetos dentro de uma coluna.



What is NoSQL?

Principais Benefícios das bases de dados de armazenamento de colunas

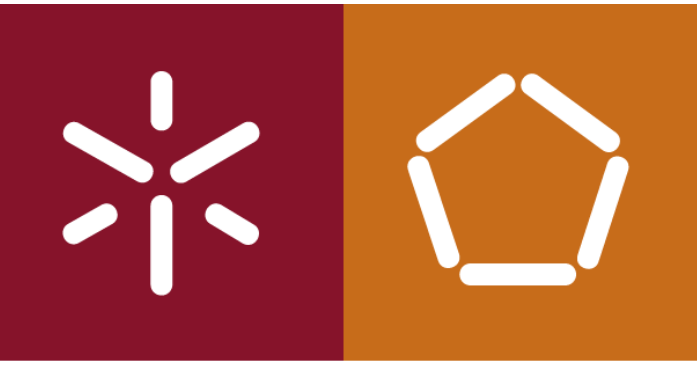
Compressão. Os armazenamentos de colunas são muito eficientes na compactação e / ou partição de dados.

Agregação. Devido à sua estrutura, as bases de dados colunares têm um bom desempenho nas consultas de agregação (como SUM, COUNT, AVG, etc.).

Escalabilidade. As bases de dados colunares são muito escaláveis. São adequados para o processamento paralelo massivo (MPP), que com disseminação de dados em arquiteturas distribuídas - geralmente milhares de máquinas.

Rápido para carregar e consultar. Armazenamentos colunares podem ser carregados rapidamente. Uma tabela de bilhões de linhas pode ser carregada em alguns segundos. Podendo ser consultada e analisada quase imediatamente.

Estes são apenas alguns dos benefícios que tornam as bases de dados colunares uma escolha popular para organizações que lidam com *big data*.



What is NoSQL?

NoSQL vs Relational

Alguns fornecedores e prestadores de serviços (como a AWS) referem as bases de dados colunares como NoSQL, quando comparadas com as bases de dados relacionais.

Muitos discordam



What is NoSQL?

Da perspectiva do utilizador, os metadados de uma base de dados colunar são equivalentes às do modelo relacional. Geridos da mesma da mesma maneira que o Oracle.

Na maioria dos casos, é 100% compatível com SQL e 100% compatível com os princípios **ACID** (ao contrário dos outros modelos NoSQL).

Os bases de dados NoSQL tendem a seguir modelos de Key Value ou repositórios de documentos.

Columnar não é se aplica a nenhum destes.

What is NoSQL?

Exemplos de *Column Store DBMSs*

- [Bigtable](#)
- [Cassandra](#)
- [HBase](#)
- [Vertica](#)
- [Druid](#)
- [Accumulo](#)
- [Hypertable](#)

What is NoSQL?

TOP 9 COLUMN-ORIENTED DATABASES

- <https://www.predictiveanalyticstoday.com/top-wide-columnar-store-databases/>

What is NoSQL?

What is a Graph Database?



What is NoSQL?

Uma base de dados gráfico é uma base de dados que usa um modelo gráfico para representar e armazenar os dados.

O modelo de base de dados de gráficos é uma alternativa ao modelo relacional.

Numa base de dados relacional, os dados são armazenados em tabelas usando uma estrutura rígida com um esquema predefinido.

Numa base de dados gráfica, não há esquema predefinido como tal. Qualquer esquema é simplesmente um reflexo dos dados que foram inseridos.

À medida que dados mais variados são inseridos, o esquema acompanha esse crescimento.

What is NoSQL?



What is NoSQL?

Os círculos azuis e verdes são nodos.
As setas representam relacionamentos.

Podemos por exemplo ver imediatamente o relacionamento que Tom Hanks tem com todos os filmes (ou seja, os círculos verdes). Mas ele também dirigiu um filme em que participou, tem dois relacionamentos com esse filme em particular.

O que é apresentado não é uma base de dados completa. São os resultados de uma consulta a uma base de dados exemplo de filmes fornecido com o sistema de de base de dados de gráficos **Neo4j**.

What is NoSQL?

As bases de dados de gráfico usam SQL?

A maioria dos sistemas de gestão de bases de dados de gráficos são sistemas NoSQL, alguns oferecem suporte para SQL, mas também suportam outras linguagens de consulta.

Na grande maioria das abordagens, o SQL não faz sentido na arquitetura do gráfico. As bases de dados de gráficos são geralmente estruturadas de maneira muito diferente das estruturas utilizadas no modelo relacional para o qual o SQL foi projetado.

What is NoSQL?

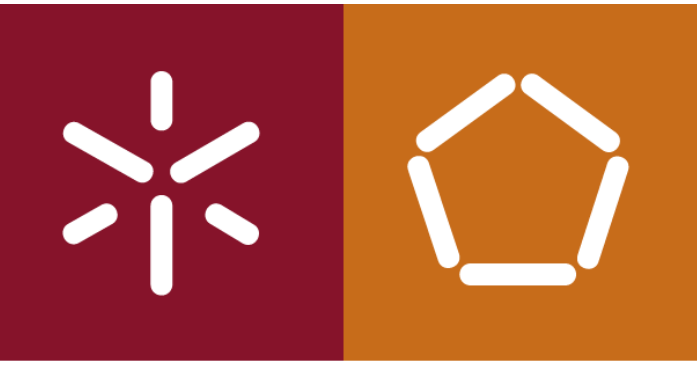
A grande parte dos sistemas de gestão de bases de dados de gráficos utilizam a sua própria linguagem de consulta.

O gráfico no exemplo anterior foi gerado a partir da seguinte consulta:

```
$MATCH (tom:Person {name: "Tom Hanks"})-[:ACTED_IN]->(tomHanksMovies) RETURN tom,tomHanksMovies
```

Este código está escrito em Cypher - a linguagem de consulta utilizada no **Neo4j**.

Alguns sistemas de bases de dados de gráficos suportam outras linguagens como *JavaScript*, *JSON*, *XQuery*, *SPARQL*, etc.



What is NoSQL?

Graph Database Management Systems (GDBMS)

A bases de dados de gráficas são criadas e geridas por um sistema de gestão de bases de dados (DBMS) projetado especificamente para bases de dados gráficas.

GDBMS (*Graph Database Management System*) ou ***Graph DBMS***.

Alguns GDBMSs utilizam mecanismos de armazenamento relacionais, enquanto outros sistemas NoSQL utilizam arquiteturas completamente diferentes para seus mecanismos de armazenamento, por exemplo, armazenamentos de valores-chave ou uma base de dados orientado a documentos.

What is NoSQL?

Muitos sistemas de bases de dados NoSQL usam **tags** ou propriedades para definir relacionamentos entre *nodos*.

Este processo ajuda no retorno de grandes quantidades de dados relacionados sem a necessidade de usar junções em várias tabelas, como seria necessário fazer no modelo relacional com o SQL.

What is NoSQL?

The screenshot displays a graph database interface with a query editor and a graph visualization. The query is:

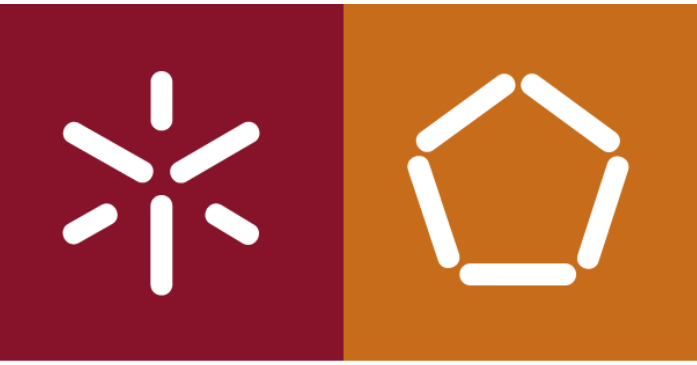
```
$ MATCH (tom:Person {name: "Tom Hanks"})-[:ACTED_IN]->
  (tomHanksMovies)-[:DIRECTED]-(director) RETURN
  tom,tomHanksMovies,director
```

The interface shows the following statistics:

- *(23) Movie(12) Person(11)
- *(26) ACTED_IN(12) DIRECTED(14)

The graph visualization shows Tom Hanks as the central node (blue circle). He is connected to 12 movies (green circles) via 'ACTED_IN' relationships and 14 other people (blue circles) via 'DIRECTED' relationships. The movies include Cloud Atlas, The Da Vinci Code, The Green Mile, Apollo 13, A League of Their Own, Cast Away, The Polar Express, Charlie Wilson's War, You've Got Mail, Sleepless, and Joe Versus the Volcano. The directors include Tom Tykwer, Ron Howard, Penny Marshall, John Patrick Stanley, and others.

At the bottom, it states: "Displaying 23 nodes, 26 relationships (completed with 26 additional relationships). AUTO-COMPLETE ON"



What is NoSQL?

A representação gráfica dos dados numa base de dados gráficos contrasta com a estrutura tabular apresentada no modelo de bases de dados relacional.

Exemplo de uma DBMS relacional

Em seguida o exemplo de uma consulta no MySQL.
Muito diferente

What is NoSQL?

```

1  SELECT
2      c.customer_id,
3      c.first_name,
4      c.last_name,
5      a.actor_id,
6      a.first_name,
7      a.last_name
8  FROM customer c
9  LEFT JOIN actor a
10 ON c.last_name = a.last_name
11 ORDER BY c.last_name;

```

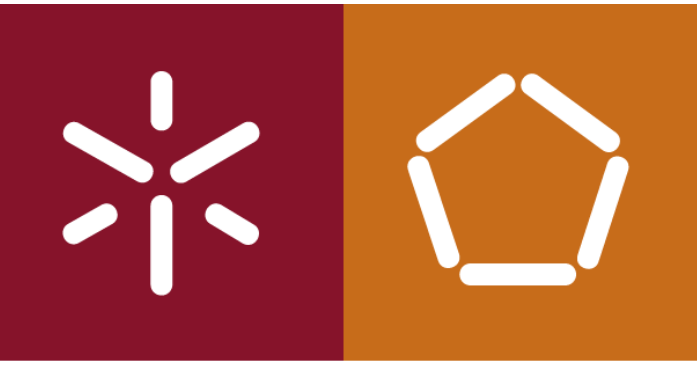
100% 22:11

Result Grid Filter Rows: Search Export:

customer_id	first_name	last_name	actor_id	first_name	last_name
504	NATHANIEL	ADAM	NULL	NULL	NULL
36	KATHLEEN	ADAMS	NULL	NULL	NULL
96	DIANA	ALEXANDER	NULL	NULL	NULL
470	GORDON	ALLARD	NULL	NULL	NULL
27	SHIRLEY	ALLEN	118	CUBA	ALLEN
27	SHIRLEY	ALLEN	145	KIM	ALLEN
27	SHIRLEY	ALLEN	194	MERYL	ALLEN
220	CHARLENE	ALVAREZ	NULL	NULL	NULL
11	LISA	ANDERSON	NULL	NULL	NULL
326	JOSE	ANDREW	NULL	NULL	NULL
183	IDA	ANDREWS	NULL	NULL	NULL

What is NoSQL?

A bases de dados de gráficas têm muitos benefícios sobre outros tipos de bases de dados, principalmente sobre as bases de dados relacionais.



What is NoSQL?

Performance

As bases de dados gráficas potenciam grandes ganhos de desempenho em relação às bases de dados relacionais, principalmente quando se trata de consultas grandes em dados relacionados.

As aplicações de bases de dados relacionais têm consultas que juntam muitas tabelas - 20, 30 ou mais. Consultas com tantas tabelas podem ser extremamente lentas, especialmente à medida que as tabelas crescem.

As consultas recursivas são também um problema específico, especialmente se tiverem elevados níveis de profundidade.



What is NoSQL?

Performance

Se esses mesmos dados forem armazenados numa base de dados gráfica, as consultas seriam mais simples e mais rápidas.

Isso porque, numa base de dados gráfica, as consultas são localizadas numa parte do gráfico. Isso significa que o tempo de execução de cada consulta é proporcional apenas ao tamanho da parte do gráfico percorrido para satisfazer as condições..

Mas tudo depende dos dados. O modelo gráfico não é necessariamente o modelo ideal para todos os casos.

Alguns dados são mais adequados à estrutura tabular das bases de dados relacionais.

No entanto podemos afirmar que o modelo gráfico é adequado para consultar grandes conjuntos de dados relacionados.

What is NoSQL?

Flexibilidade

Como não há necessidade de definir um esquema, temos total flexibilidade sobre o crescimento da base de dados.

No modelo relacional, precisamos de mapear os requisitos em detalhe antes de criar a base de dados.

Precisa de prever qualquer possível mudança nos requisitos e criar uma solução que atenda a todos os cenários futuros possíveis.

Isto nem sempre é possível.

Se os requisitos de negócios aumentarem ou mudarem significativamente, a estrutura da base de dados também poderá precisar mudar significativamente.

Podemos ter de a refazer completamente.

What is NoSQL?

Flexibilidade

Com o modelo gráfico, à medida que a empresa cresce, pode adicionar novos tipos de relacionamentos.

Novos nodos e novos subgráficos sem perturbar as consultas e as funcionalidades da camada aplicacional.

Esta facilidade tem também implicações na facilidade de desenvolvimento incremental de aplicações.

What is NoSQL?

Exemplos de aplicação

- Social networks
- Realtime product recommendations
- Network diagrams
- Fraud detection
- Access management
- Graph based search of digital assets
- Master data management

What is NoSQL?

Exemplos de *Graph Databases*

Neo4j, Blazegraph, and OrientDB.

What is NoSQL?

The (2)7 Best Graph Databases to Consider for 2020

<https://solutionsreview.com/data-management/the-best-graph-databases-to-consider-in-2020/>

<https://www.predictiveanalyticstoday.com/top-graph-databases/>

What is NoSQL?

- ◆ Connolly, T., Begg, C., **Database Systems, A Practical Approach to Design, Implementation, and Management** , Addison-Wesley, 6a Edição, 2014.
- ◆ Luc Perkins, Eric Redmond, Jim Wilson, **Seven Databases in Seven Weeks - A Guide to Modern Databases and the NoSQL Movement**, Pragmatic Bookshelf, 2018.