

Alfabetização de Dados

Compreenda, organize e interprete os dados da sua organização

Big Data e NoSQL

Capítulo 6

Agenda

- ❑ Big Data
 - ❑ MapReduce
- ❑ Banco de Dados NoSQL:
 - ❑ Banco de Dados Modelo Chave-Valor
 - ❑ Banco de Dados Modelo Documentos
 - ❑ Banco de Dados Modelo Família de Colunas
 - ❑ Banco de Dados Modelo Grafos



Big Data

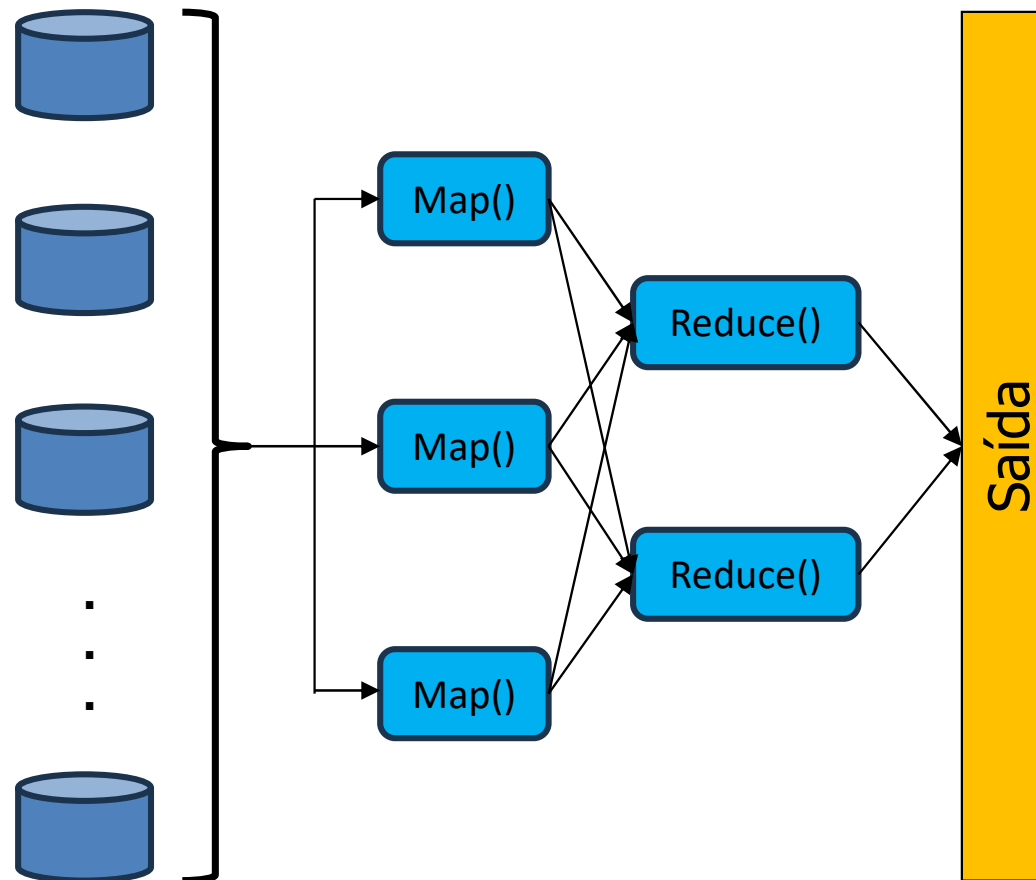
O big data é caracterizado por um grande `volume` de dados, adquiridos e processados em uma `velocidade` apropriada para não se tornarem obsoletos e disponibilizados em uma grande gama de `variedade` de formatos.

Big Data

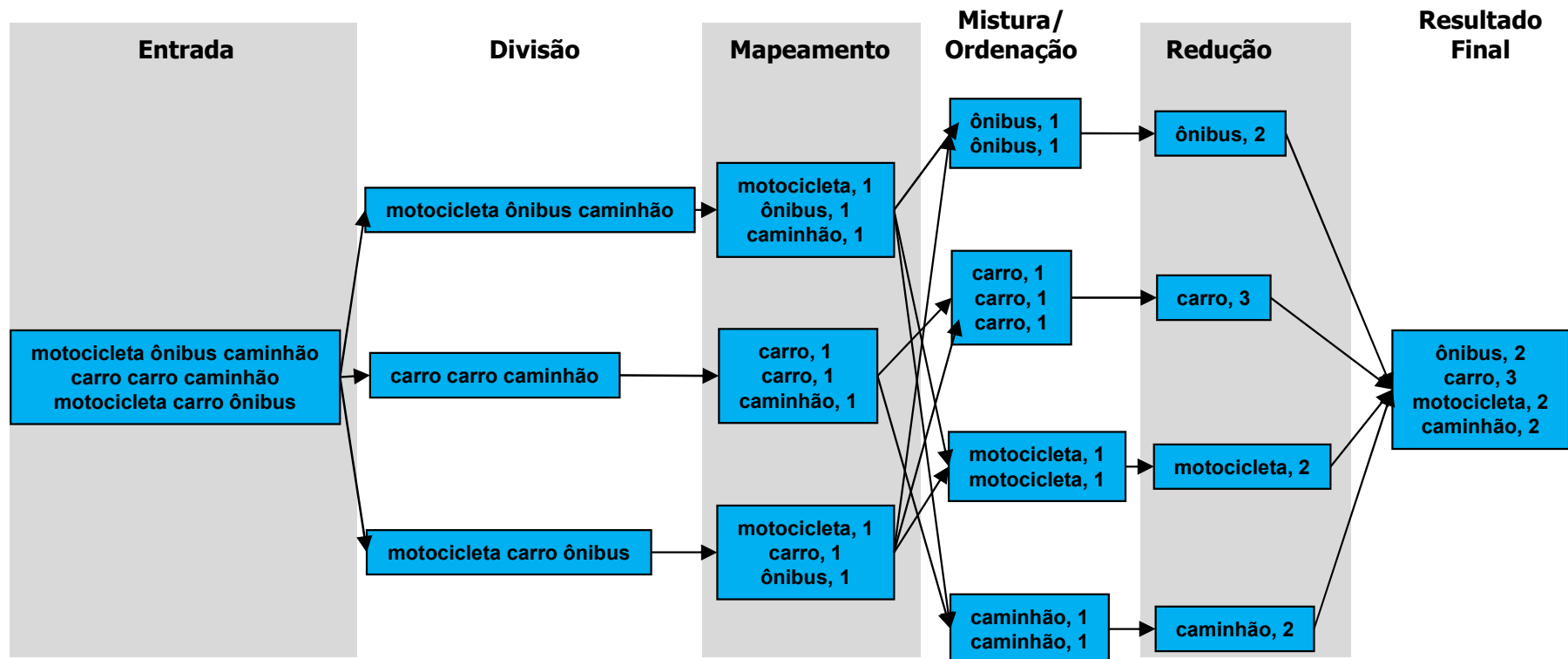
- ❑ O MapReduce, pode ser dividido em dois estágios principais:
 - ❑ **Mapeamento:** o nó mestre divide os dados em vários subconjuntos menores, um nó trabalhador processa um subconjunto de dados menor sob o controle de um rastreador de trabalho e armazena o resultado no sistema de arquivos local, onde um redutor será capaz de acessá-lo.
 - ❑ **Redução:** esta etapa analisa e reúne os dados de entrada a partir das etapas de mapeamento. Pode haver múltiplas tarefas de redução para paralelizar a agregação, e essas tarefas são executadas nos nós trabalhadores sob o controle do rastreador de trabalho.

MapReduce

Big Data



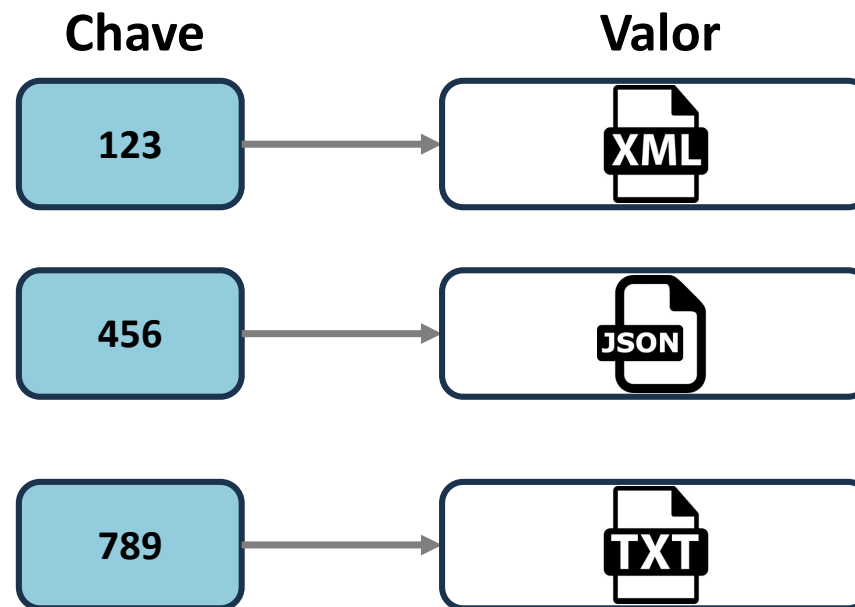
MapReduce



Bancos de Dados NoSQL

- ❑ As principais características dos SGBDs NoSQL são:
 - ❑ não utilizam o SQL (possuem linguagens semelhantes);
 - ❑ geralmente são projetos de código aberto;
 - ❑ a maioria dos BDs é orientada pela necessidade de execução distribuída;
 - ❑ oferecem uma gama de opções para consistência e distribuição; e
 - ❑ não possuem esquemas rígidos.

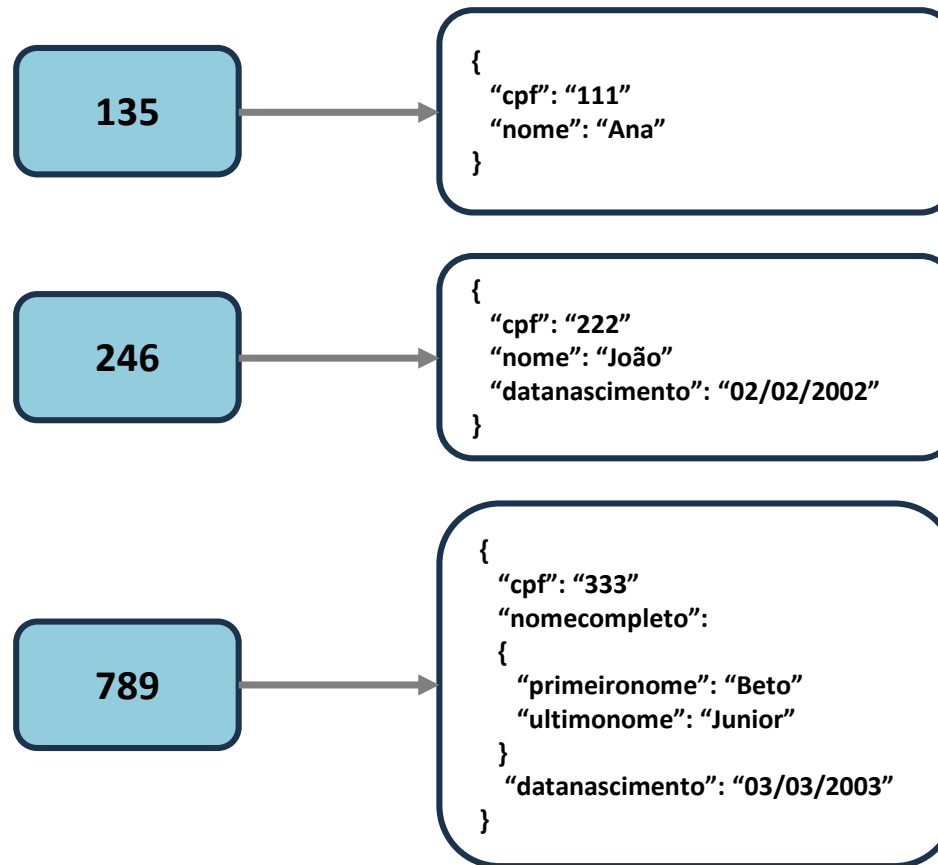
Modelo Chave-Valor



Modelo Documentos

Identificador

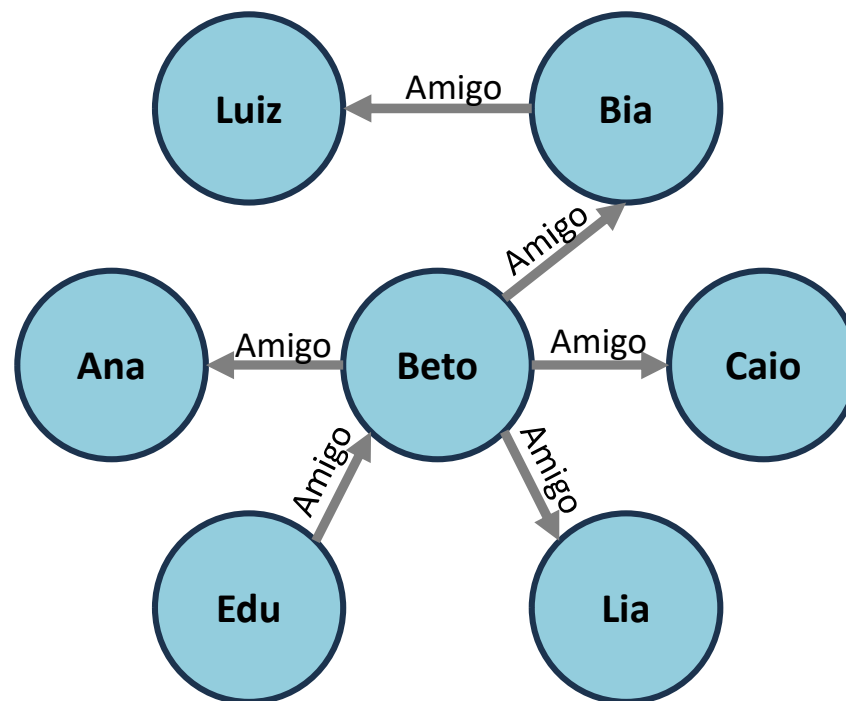
Documento

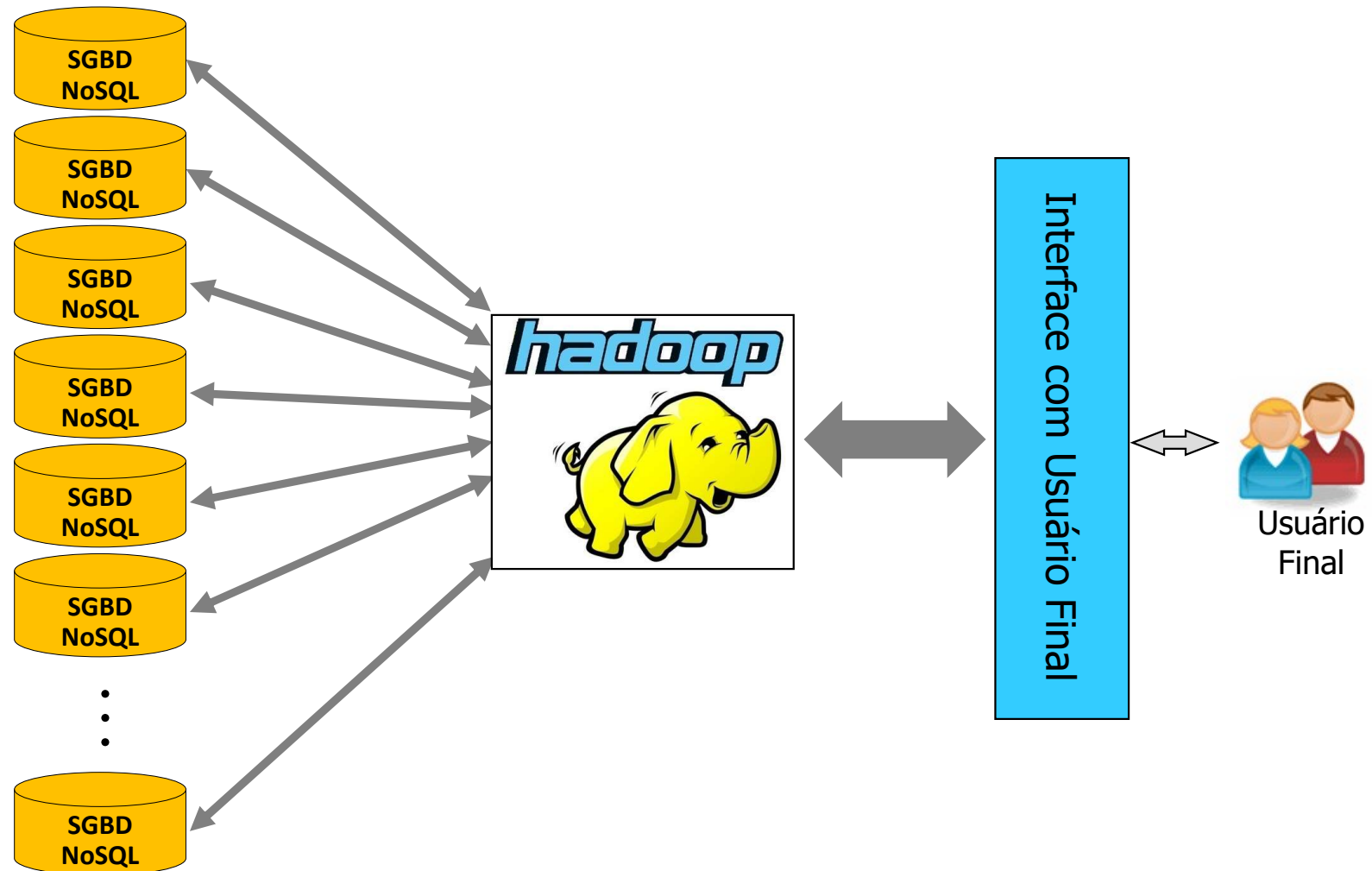


Modelo Família de Colunas

CPF	Nome	Sobrenome	Nascimento	Pais
111	Ana	Silva	01/01/2001	Brasil
222	Beto	Lima		
333	Caio	Abreu	03/02/2003	
444	Lia	Correia		

Modelo Grafos





Fim