

UFSC-CTC-INE

INE 5600 – Bancos de Dados III

BD NoSQL: Modelagem

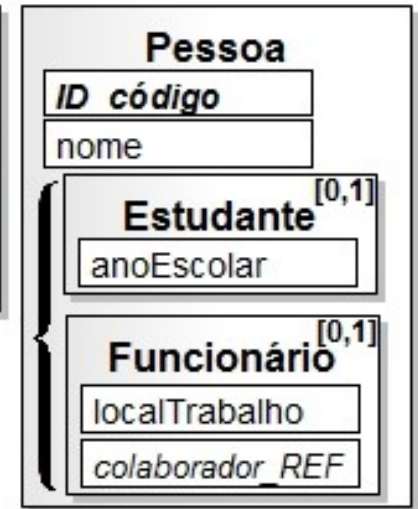
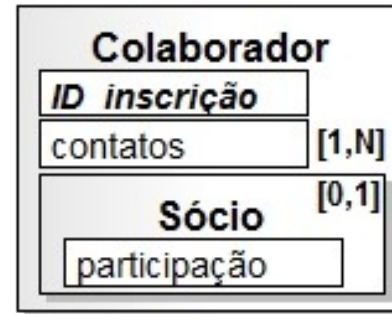
Projeto de BDs NoSQL

- Tema de pesquisa atual
- Modelagem de BDs NoSQL é um tópico pouco explorado na academia e na indústria
 - nem todo BD NoSQL exige um esquema para os dados, porém...
 - a existência de um esquema requer decisões de projeto para evitar desempenho ruim no acesso à *Big Data*
- Conceito de **Agregado** (Sadalage & Fowler, 2012)
 - base para a modelagem lógica de BDs de documento, colunar e chave-valor
 - representação de dados relacionados de forma (preferencialmente) aninhada

Modelo Lógico de Agregados

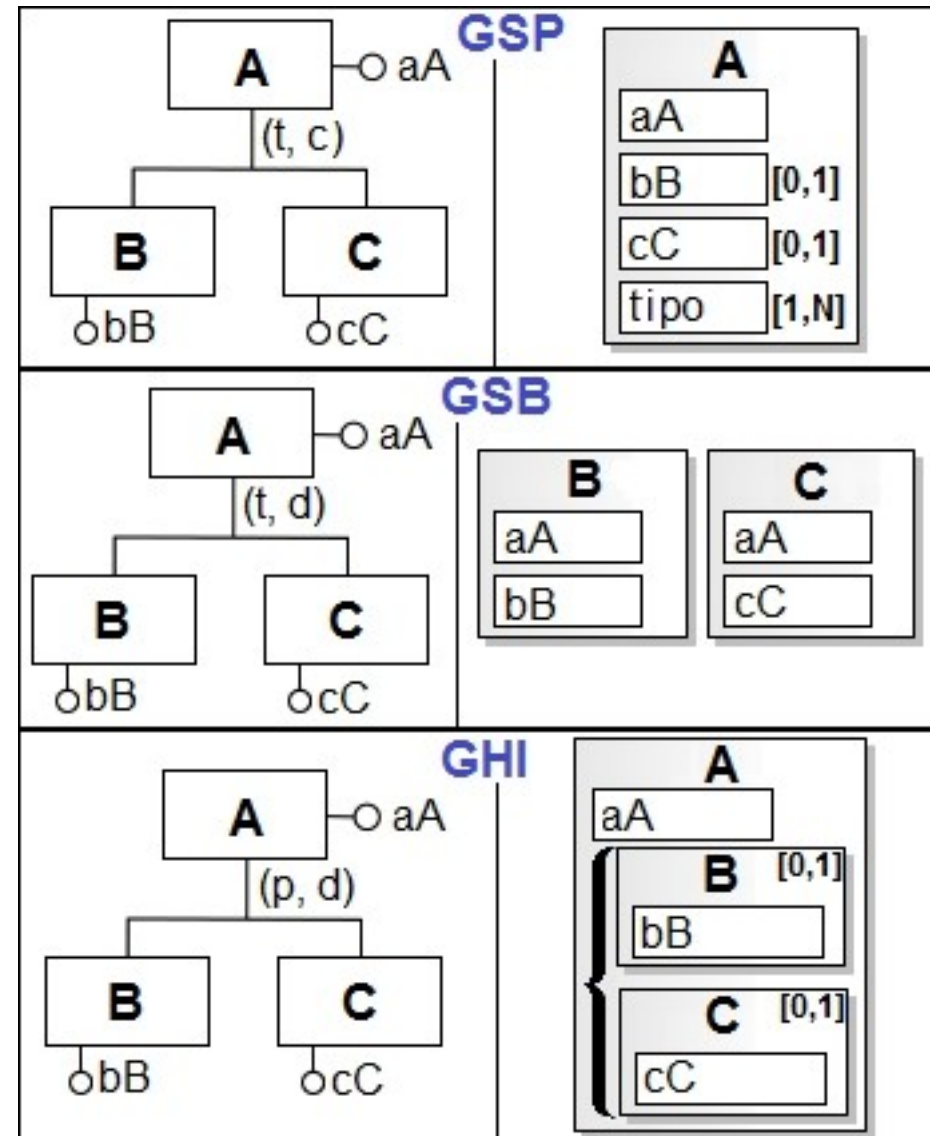
- Conceitos básicos

- coleções, blocos e atributos
- um esquema de BD possui uma ou mais coleções
- uma coleção é composta por um bloco raiz (unidade de acesso)
- um bloco raiz é composto por um ID e um conjunto não-vazio de atributos e/ou blocos aninhados (mono ou multivalorados)
- tipos de relacionamento
 - Hierárquicos (blocos aninhados) (pode haver disjunção)
 - Referência (atributo)



Mapeamento Conceitual-Lógico

- Três alternativas para a conversão de generalizações
 - GSP** – Ênfase na **Superclasse**
 - GSB** – Ênfase nas **Subclasses**
 - GHI** – Ênfase na **Hierarquia**
- Decisão por uma das alternativas deve considerar
 - tamanho da porção do esquema gerado
 - prioridades de acesso (*workload*)
 - restrições (*totalidade* e *disjunção*)



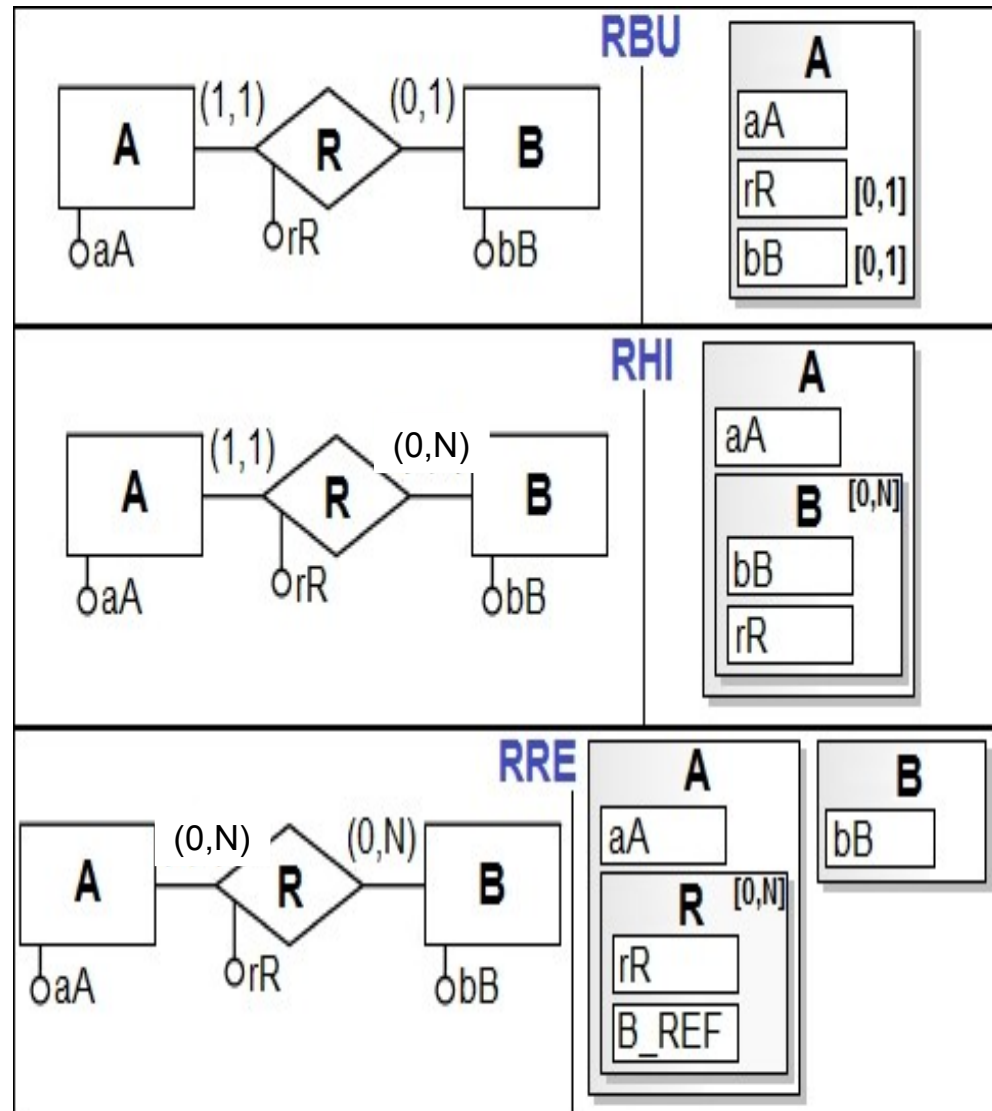
Mapeamento Conceitual-Lógico

- Três alternativas para a conversão de relacionamentos

- RBU** – Modelado por **Bloco Único**
- RHI** – Modelado por **Hierarquia**
- RRE** – Modelado por **Referências**

- Decisão por uma das alternativas deve considerar as cardinalidades do relacionamento e as prioridades de acesso

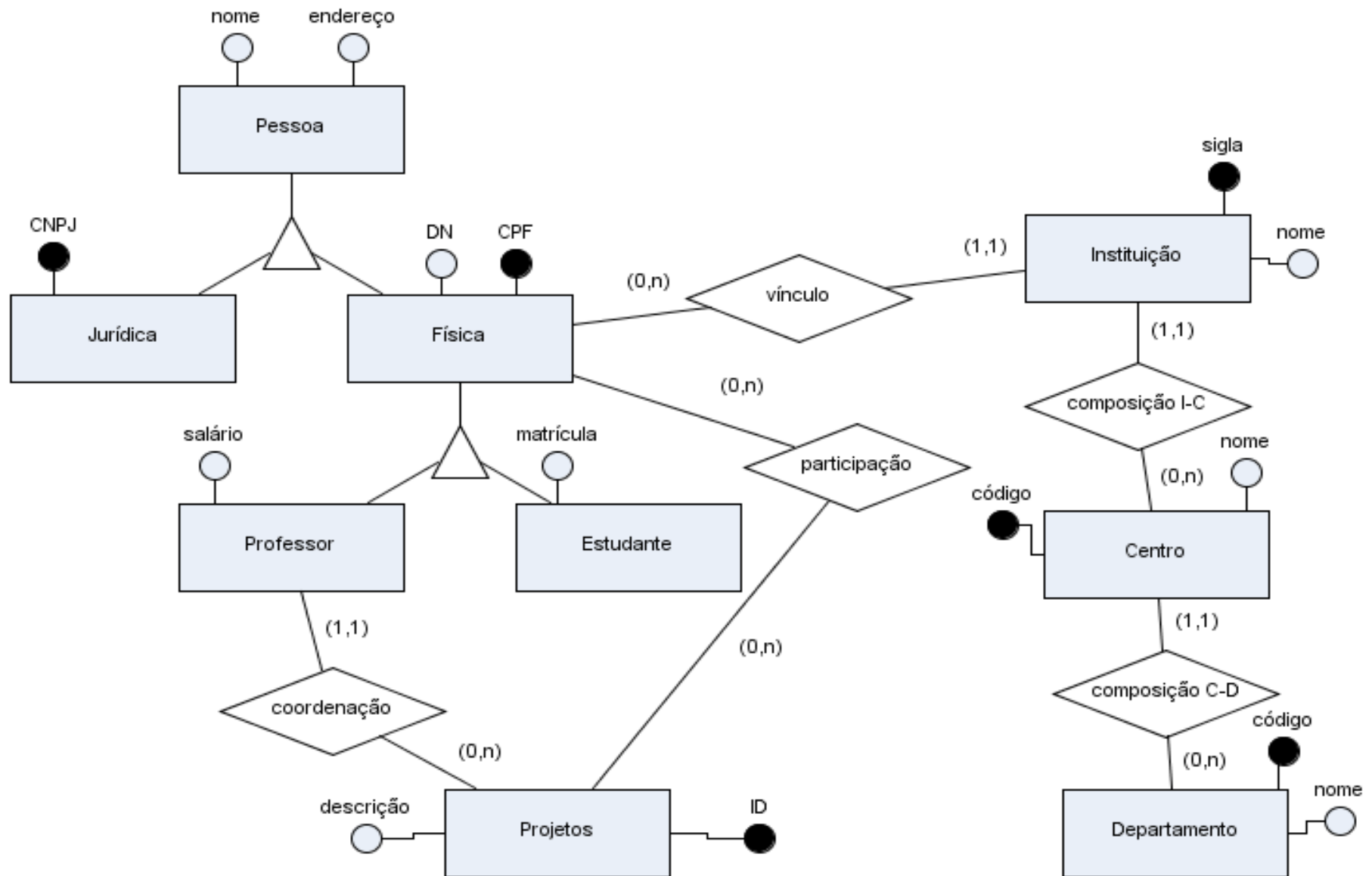
- RBU : (1:1)**
- RHI : (1:N) e (1:1)**
- RRE : (M:N), (1:N) e (1:1)**



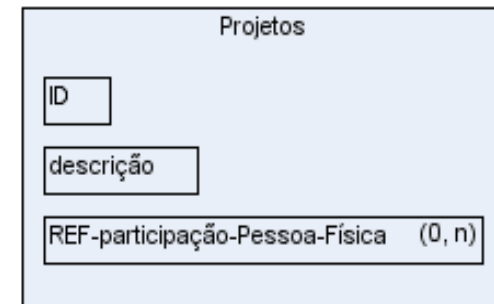
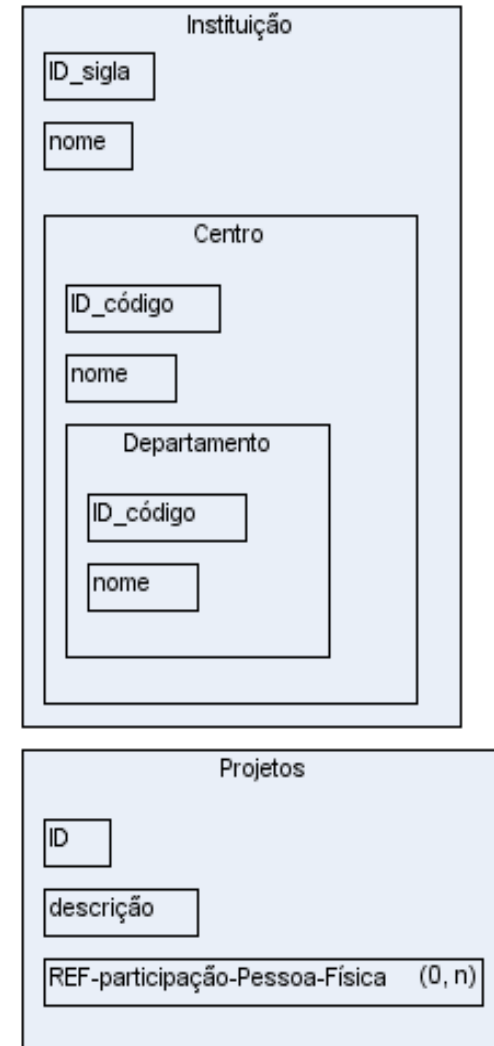
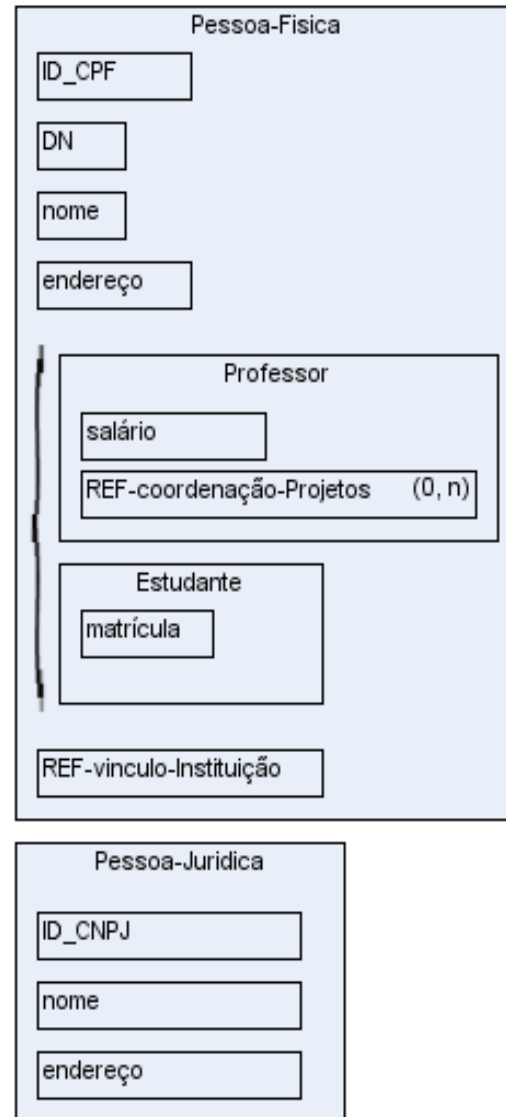
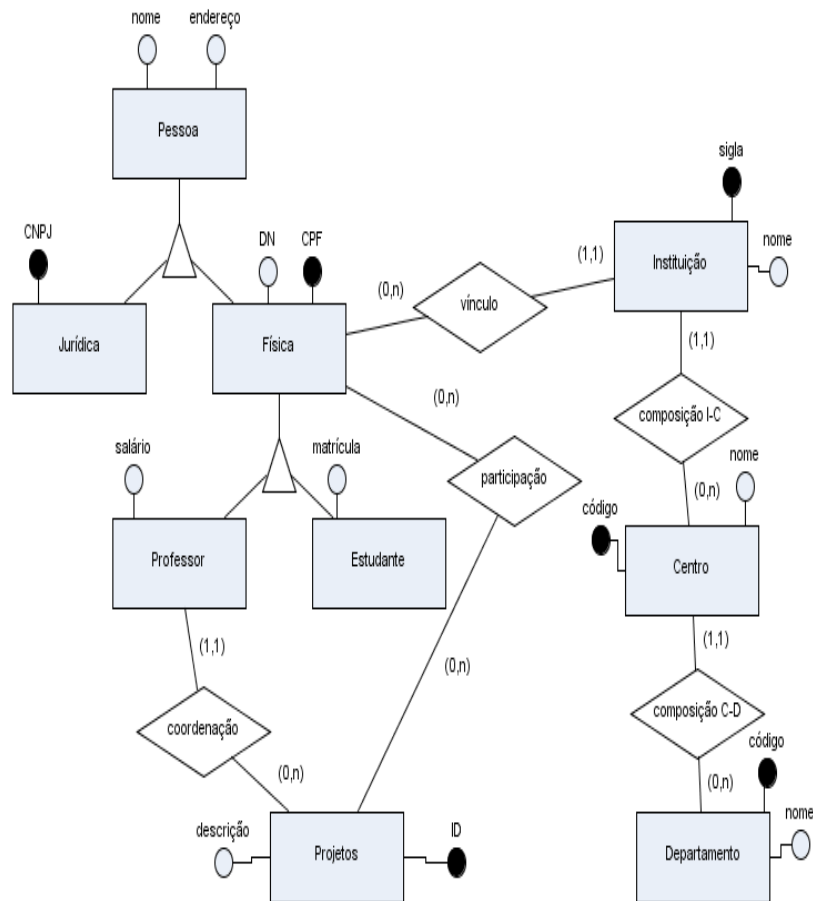
Processo de Projeto Lógico

- Conversão de Hierarquias
 - prioridade para a geração da **menor porção de esquema**
 - estratégia **bottom-up** de conversão para hierarquias de vários níveis
 - regra **GSB** não se aplica a generalizações parciais (p)
- Conversão de Relacionamentos
 - prioridade para a conversão de **cadeias mais longas de relacionamentos com cardinalidade 1-1 ou 1-N**
 - promove um maior nível de aninhamento de dados

Exemplo de Modelagem Conceitual



Possível Modelagem Lógica NoSQL

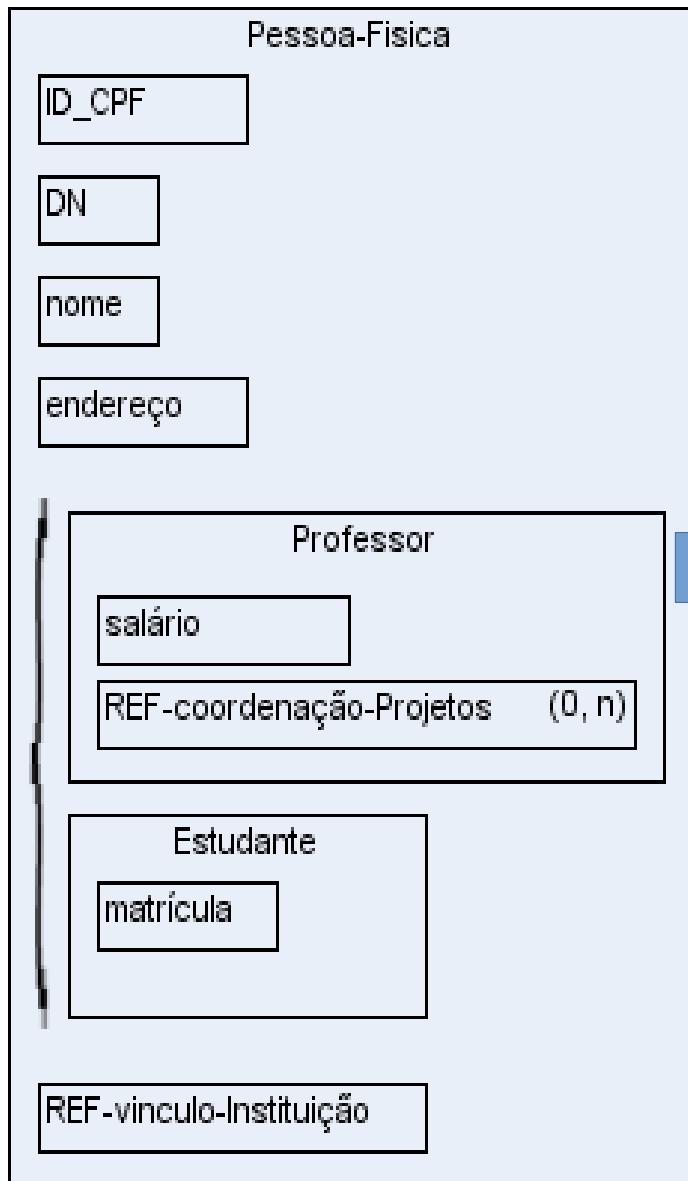


Modelagem Física NoSQL – BD Chave-Valor

- Coleção \Rightarrow Par Chave-Valor
- Regras de Mapeamento
 - ID da coleção \Rightarrow Chave
 - Conteúdo da coleção \Rightarrow Valor *

* conteúdos estruturados devem ser desaninhados e serializados

Exemplo Modelagem Física BD NoSQL Chave-Valor



BD Chave-Valor

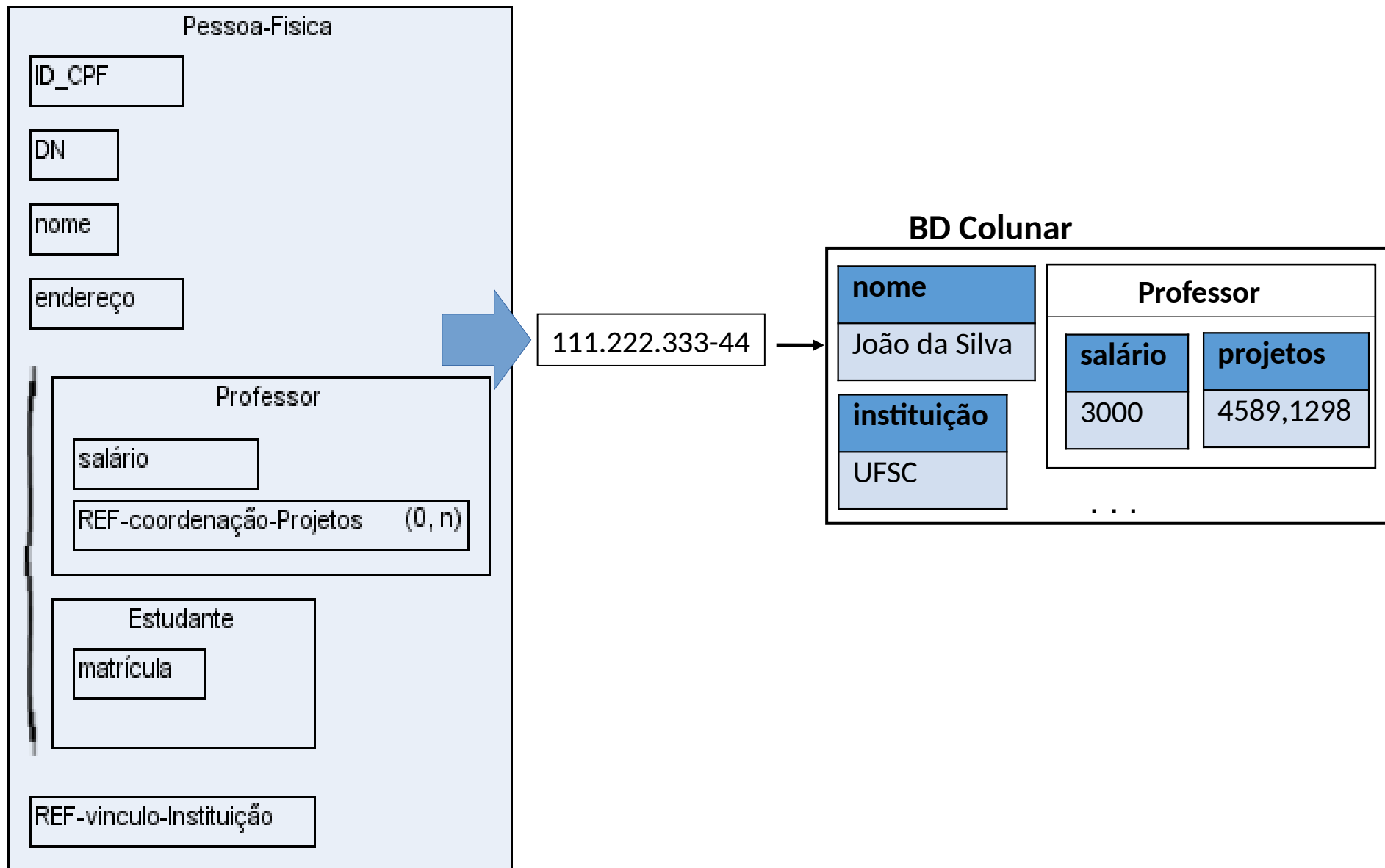
Chave: 111.222.333-44

Valor: "nome: João da Silva; ...; salário: 3000; projetos: 4589,1298; instituição: UFSC"

Modelagem Física NoSQL – BD Colunar

- Coleção \Rightarrow Família de Colunas
- Regras de Mapeamento
 - ID da coleção \Rightarrow Chave da família de colunas
 - Atributo simples \Rightarrow Coluna
 - Atributo multivalorado \Rightarrow Coluna multivalorada ou coluna com conteúdo serializado
 - Atributo de referência \Rightarrow Chave da coleção referenciada
 - Bloco \Rightarrow Supercoluna

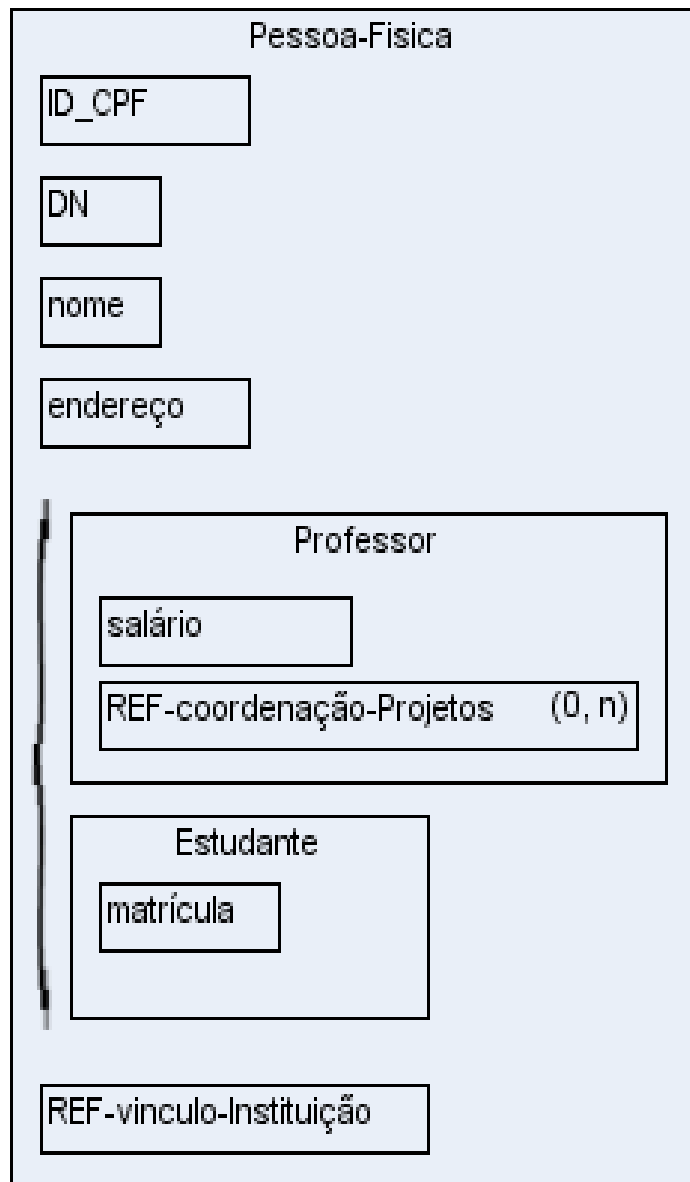
Exemplo Modelagem Física BD NoSQL Colunar



Modelagem Física NoSQL – BD Documento

- Coleção \Rightarrow Documento
- Regras de Mapeamento
 - ID da coleção \Rightarrow Chave do documento
 - Atributo simples \Rightarrow Atributo simples
 - Atributo multivalorado \Rightarrow Atributo do tipo Lista
 - Atributo de referência \Rightarrow Chave da coleção referenciada
 - Bloco \Rightarrow Atributo do tipo Objeto

Exemplo Modelagem Física BD NoSQL Documento



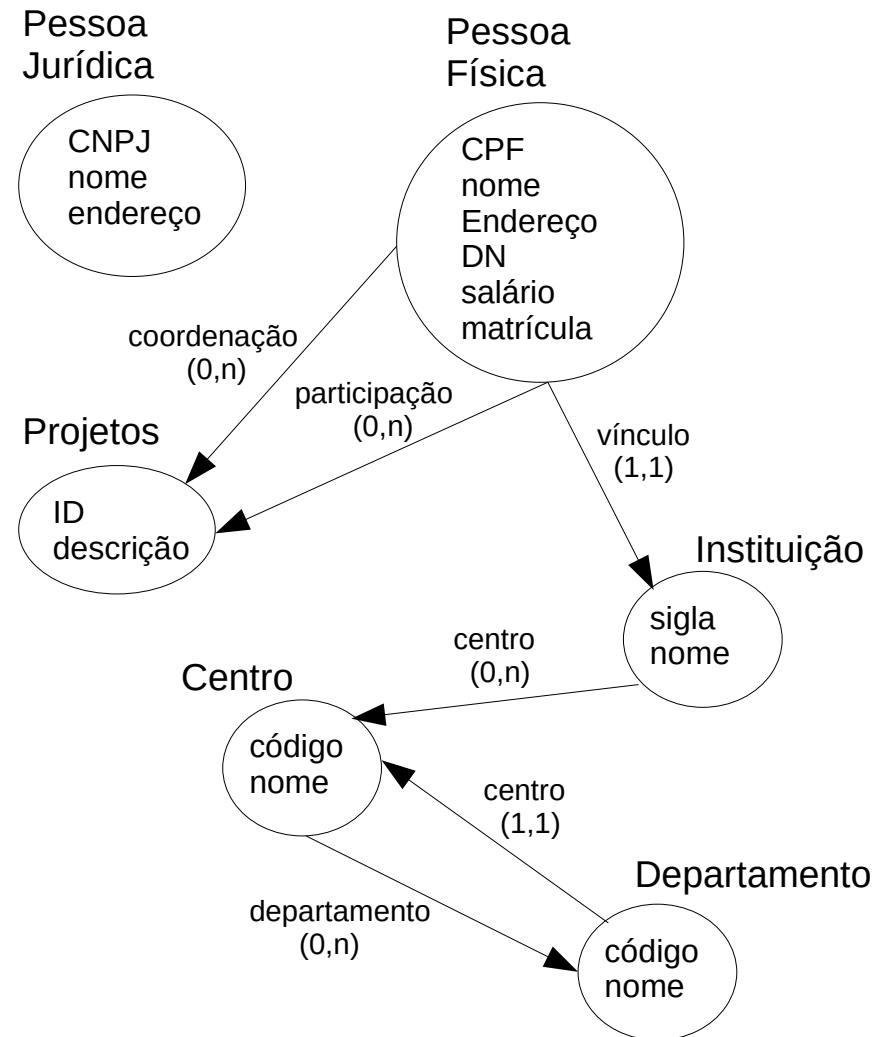
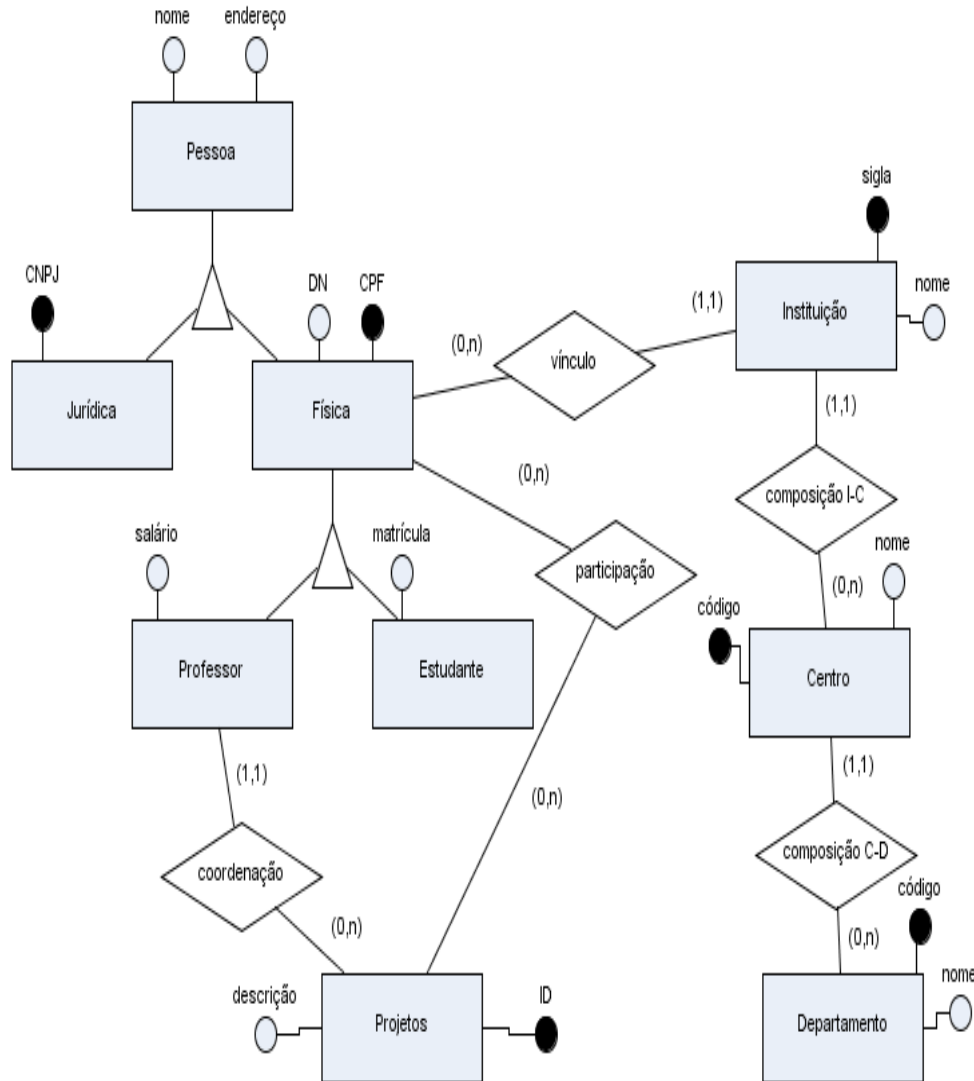
BD Documento

```
{ "_id": "111.222.333-44",  
  "nome": "João da Silva", ...  
  "Professor": {  
    "salário": "3000",  
    "projetos": [ "4589", "1298" ],  
    "instituição": "UFSC" }  
}
```

BD NoSQL Grafo

- Não segue uma abordagem orientada a agregados
 - Ênfase nos relacionamentos entre quaisquer tipos de dados e não em uma organização hierárquica deles
- Neste caso, basicamente entidades tornam-se nodos e relacionamentos tornam-se arestas
 - Decisão por arestas unidirecionais ou bidirecionais depende do *workload*
 - Pode-se fundir entidades fracamente relacionadas à entidades mais conectadas

Exemplo Modelagem Lógica BD NoSQL Grafo



Atividade 4 – Projeto NoSQL

Proponha um projeto lógico NoSQL baseado em agregados e um projeto lógico para BD NoSQL grafo, para o domínio de um **Zoológico** mostrado abaixo.

