

# Bases de données NoSQL

Compte rendu

---

## Rendu sur CouchDB et MapReduce

---

Enseignant:

Samir Youcef

Auteurs:

Ghada Zeineb MALEK (INFO3)

Année : 2023 - 2024

# Table des Matières

<b>1</b>	<b>Présentation de CouchDB</b>	<b>2</b>
1.1	Installation de CouchDB . . . . .	2
1.2	Manipulation des données avec CouchDB . . . . .	2
<b>2</b>	<b>Fonction MapReduce</b>	<b>3</b>
2.1	Présentation . . . . .	3
2.2	Exemples d'utilisation . . . . .	3
2.2.1	Calcul du nombre de films par année . . . . .	3
2.2.2	Calcul du nombre de films par acteur . . . . .	4

# Introduction

CouchDB est un gestionnaire de base de données documentaire de type NoSQL. Ce compte rendu présente son installation, ses principales fonctionnalités via les API REST, ainsi que l'utilisation de la fonction **MapReduce** pour effectuer des traitements de données massifs.

## 1 Présentation de CouchDB

CouchDB est un système de base de données NoSQL orienté documents. Il propose des API REST pour l'interaction avec les données.

- **Facilité d'utilisation** : CouchDB est simple à installer et utiliser.
- **Client graphique** : Accessible via le port 5984.
- **Client ligne de commande** : Utilisation de `curl` pour communiquer avec l'API REST.

### 1.1 Installation de CouchDB

Deux méthodes principales permettent d'installer CouchDB :

1. **Installation native** (sur un OS).
2. **Utilisation de Docker** :

- Exécution avec la commande suivante :

```
docker run -d --name couchdbdemo \\  
-e COUCHDB_USER=youcef -e COUCHDB_PASSWORD=samir \\  
-p 5984:5984 couchdb
```

- Accès au client graphique : `http://localhost:5984/_utils`.
- Connexion via ligne de commande avec `curl` :

```
curl -X GET http://youcef:samir@localhost:5984
```

### 1.2 Manipulation des données avec CouchDB

**Insertion d'un document :**

- Utilisation de la méthode **PUT** pour créer une base de données :

```
curl -X PUT http://youcef:samir@localhost:5984/films
```

- **Insertion d'un document** : CouchDB génère un identifiant unique s'il n'est pas spécifié.

**Exemple d'insertion d'une collection de films :**

```
curl -X POST http://youcef:samir@localhost:5984/films \\  
-H "Content-Type: application/json" \\  
-d '{"title": "Inception", "year": 2010, "actors": [{"first_name": "Leonardo", "last_name": "DiCaprio"]}'
```

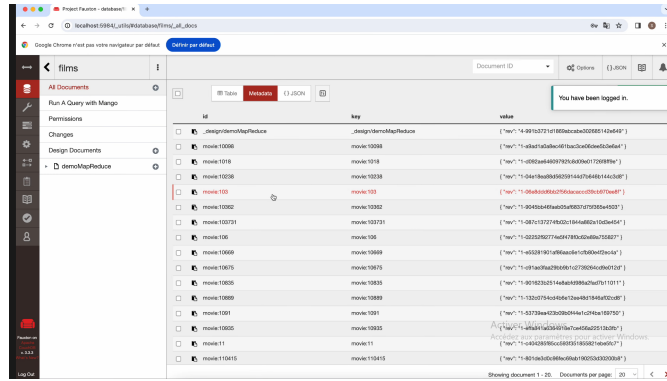


Figure 1: Interface graphique

## 2 Fonction MapReduce

### 2.1 Présentation

MapReduce est une technique utilisée pour traiter des volumes de données importants en parallèle. CouchDB implémente MapReduce à travers deux fonctions principales : **Map** et **Reduce**.

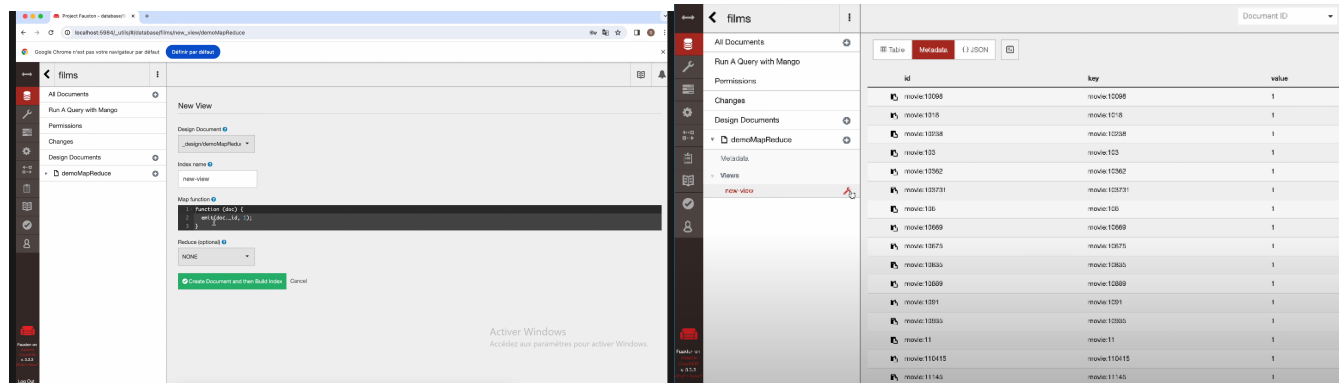
- **Fonction Map** : Appliquée à chaque document, elle permet de transformer les données et de définir des clés de regroupement.
- **Fonction Reduce** : Effectue une agrégation des clés intermédiaires.

### 2.2 Exemples d'utilisation

#### 2.2.1 Calcul du nombre de films par année

Fonction Map :

```
function (doc) {
    emit(doc.year, doc.title);
}
```



(a) Fonction map

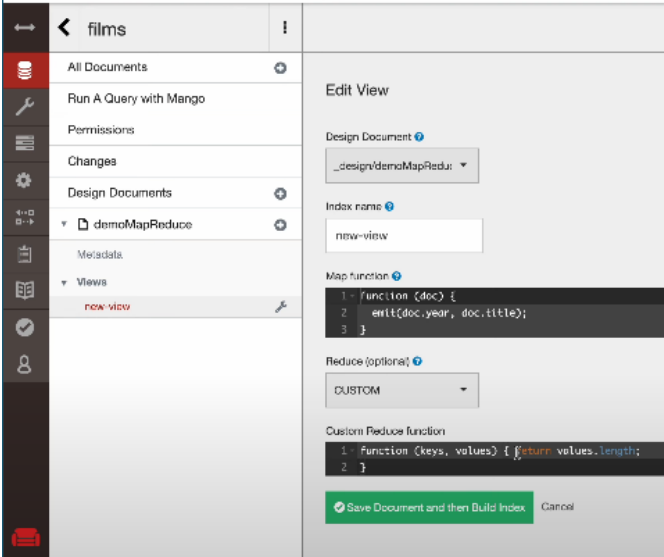
(b) Resultat fonction map

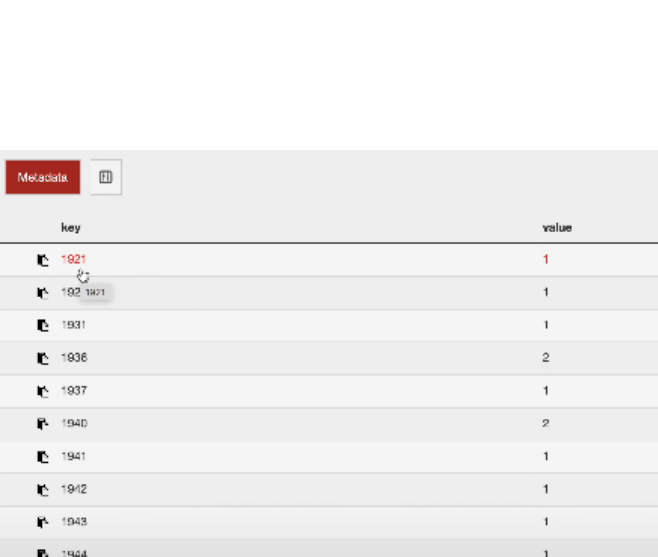
Chaque *map* traite un sous-ensemble des documents pour extraire l'année et le titre du film.

## Fonction Reduce :

```
function (keys, values) {
  return values.length;
}
```

Cette fonction permet d'agréger le nombre de films par année.





key	value
1921	1
1921-1921	1
1931	1
1936	2
1937	1
1940	2
1941	1
1942	1
1943	1
1944	1

(a) Fonction map

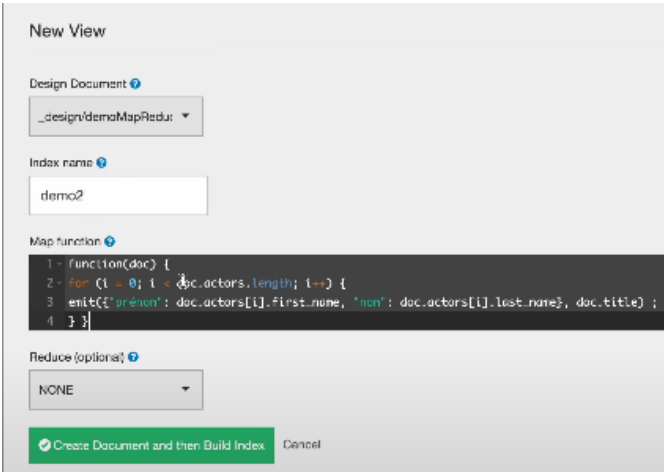
(b) Resultat fonction map

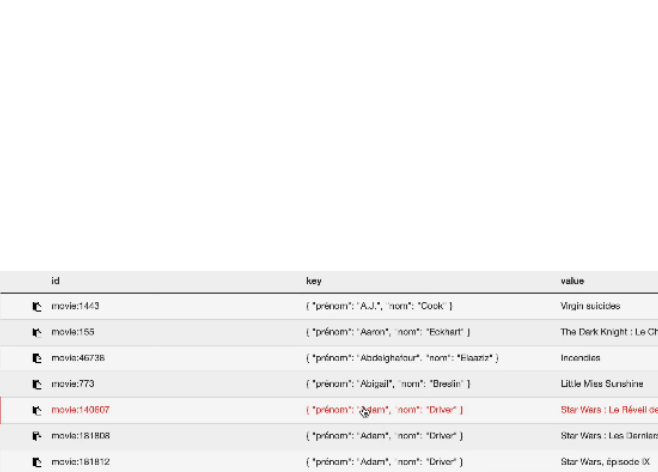
## 2.2.2 Calcul du nombre de films par acteur

### Fonction Map :

```
function (doc) {
  for (i = 0; i < doc.actors.length; i++) {
    emit({"prenom": doc.actors[i].first_name, "nom": doc.actors[i].last_name}, doc.title);
  }
}
```

Cette fonction parcourt le tableau d'acteurs pour extraire le prénom, le nom, et le titre du film.





id	key	value
movie:1443	{"prenom": "A.J.", "nom": "Cook"}	Virgin suicides
movie:155	{"prenom": "Aaron", "nom": "Eckhart"}	The Dark Knight : Le Chevalier sombre
movie:46736	{"prenom": "Abdourahmane", "nom": "Eliaass"}	Incendies
movie:773	{"prenom": "Abigail", "nom": "Breslin"}	Little Miss Sunshine
movie:140607	{"prenom": "Adam", "nom": "Driver"}	Star Wars : Le Réveil de la Force
movie:181808	{"prenom": "Adam", "nom": "Driver"}	Star Wars : Les Derniers Jedi
movie:181812	{"prenom": "Adam", "nom": "Driver"}	Star Wars, épisode IX

(a) Fonction map

(b) Resultat fonction map

### **Fonction Reduce :**

```
function (keys, values) {  
    return values.length;  
}
```

## **Conclusion**

CouchDB est un outil puissant pour la gestion de bases de données documentaires, offrant des capacités de traitement massives via MapReduce. Sa simplicité d'utilisation et ses fonctionnalités de parallélisation en font un choix adapté pour les systèmes distribués modernes.