

Bases de données NoSQL

Rapport de TP



Auteur : IHSSANE ER-RAMI
Encadrant : M. YUCEF SAMIR

Table des matières

1	Introduction	3
1.1	Définition de CouchDB	3
1.2	Installation de CouchDB	3
1.2.1	Docker	3
1.3	Vérification du serveur CouchDB	3
2	Manipulation des bases de données	3
2.1	Création d'une bdd	3
2.2	Insertion de documents	3
2.3	Récupération d'un document	3
3	MapReduce avec CouchDB	4
3.1	Définition	4
3.2	Nombre de films par année	4
3.2.1	map	4
3.2.2	reduce	4
4	Conclusion	4

1 Introduction

1.1 Définition de CouchDB

CouchDB est une base de données NoSQL permettant le stockage de données sous forme de documents JSON. Ces documents sont indexés et traités via des vues exploitant le paradigme MapReduce. Conçu pour être flexible, CouchDB facilite le traitement distribué de vastes volumes de données.

Dans le cadre de ce TP, CouchDB sera utilisé pour stocker des matrices représentant les liens entre différentes pages web et effectuer des calculs à l'aide de MapReduce. Ses principales caractéristiques sont les suivantes :

-Installation et utilisation simples. -Open Source, sous licence Apache. -Prise en charge d'une API REST, permettant : -GET : récupération de ressources, -PUT : création de ressources, -POST : envoi de données, -DELETE : suppression de ressources.

1.2 Installation de CouchDB

1.2.1 Docker

utilisation de la commande suivante :

```
docker run -d --name couchDbDemo -e COUCHDB_USER=abc -e COUCHDB_PASSWORD=123 -p 5984:5984 couchdb
```

L'interface graphique est accessible via :

http://localhost:5984/_utils

1.3 Vérification du serveur CouchDB

On utilise la commande suivante :

```
curl -X GET http://abc:123@localhost:5984/
```

2 Manipulation des bases de données

2.1 Création d'une bdd

```
curl -X PUT http://abc:123@localhost:5984/films
```

2.2 Insertion de documents

pour un doc simple :

```
curl -X PUT http://abc:123@localhost:5984/films/doc1 -d '{"titre":"Inception", "annee":2010}'
```

pour multiples docs ou insertion en masse :

```
curl -X POST http://abc:123@localhost:5984/films/_bulk_docs -d @films_couchdb.json -H "Content-Type: appli"
```

2.3 Récupération d'un document

```
curl -X GET http://abc:123@localhost:5984/films/movie:1001
```

3 MapReduce avec CouchDB

3.1 Définition

MapReduce est un modèle de programmation utilisé pour traiter efficacement de grands ensembles de données en parallèle. Il repose sur deux étapes :

Map : chaque document est traité individuellement et produit une paire clé-valeur. Reduce : les paires clé-valeur sont agrégées pour générer un résultat final.

3.2 Nombre de films par année

3.2.1 map

```
function (doc) {  
  emit(doc.annee, doc.titre);  
}
```

a pour résultat :

```
{"key":2010, "value":"Inception"}  
{"key":2014, "value":"Interstellar"}
```

3.2.2 reduce

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

4 Conclusion

CouchDB, avec son approche flexible et son intégration de MapReduce, est une solution idéale pour la gestion de bases de données documentaires et le traitement de grandes quantités de données distribuées.