

TP1 CouchDB (3 heures)

1. Préalable

Vous travaillerez sur le cluster distribué CouchDB (version 3.3) installé sur les serveurs de la DSIN, constitué de 4 nœuds, et accessible aux adresses de ces quatre nœuds, qui sont :

```
prodpeda-couchdb-1.infra.umontpellier.fr:5984,  
prodpeda-couchdb-2.infra.umontpellier.fr:5984,  
prodpeda-couchdb-3.infra.umontpellier.fr:5984  
et prodpeda-couchdb-4.infra.umontpellier.fr:5984.
```

2. Création et alimentation de la base de données

Vous définirez au préalable une variable d'environnement nommée COUCH3 dans votre .bashrc pour vous simplifier l'accès au cluster en ligne de commande. Les informations concernant la chaîne de connexion sont données dans un fichier textuel à part (nommé DansBashrc_24.txt).

Il est parfois utile de ne pas mobiliser le cache du serveur Web lors de requêtes http émises via le client curl (sous peine d'avoir une réponse vide). Utilisez alors la directive -H "Cache-Control: no-cache", voir par exemple la commande : curl -H "Cache-Control: no-cache" ...

- Vous définirez une base de données nommée **occitanie** en la préfixant avec votre nom de famille (tout en minuscules). Réalisez cette opération en ligne de commandes avec curl ;
- Vous exploiterez le mode **"ajout par lots"** via curl pour insérer les documents des fichiers herault.json, gard.json, aveyron.json, hauteGaronne.json et regions_partiel.json dans la base de données.

Le contexte applicatif est le même que celui abordé pour Neo4J, à savoir des informations administratives sur des communes françaises. Des exemples de documents json extraits des fichiers de commune et de région sont proposés ici. Vous noterez le champ **type** qui est une tentative de définition d'un schéma.

```
{  
  "_id": "34100",  
  "codeInsee": "34100",  
  "longitude": "2.8925",  
  "latitude": "43.4858",  
  "type": "commune",  
  "nom": "FERRIERES-POUSSAROU",  
  "dep": "34",  
  "old_reg": "91",  
  "populations": [  
    {  
      "pop_1975": 38.0961632716736  
    },  
    {  
      "pop_1985": 35.8391614288578  
    },  
    {  
      "pop_1995": 45.2742630644936  
    },  
    {  
      "pop_2005": 56.941818739832  
    },  
  ],  
}
```

```
{
  "pop_2010": 60.6582030010204
}
```

Listing 1 – Exemple de commune

```
{
  "_id": "91",
  "reg": "91",
  "chef_lieu_reg": "34172",
  "type": "old_region",
  "new_reg": "occitanie",
  "nom_reg": "LANGUEDOC-ROUSSILLON",
  "departements": [
    {
      "dep": "34",
      "nom_dep": "HERAULT",
      "chef_lieu_dep": "34172"
    },
    {
      "dep": "30",
      "nom_dep": "GARD",
      "chef_lieu_dep": "30189"
    },
    {
      "dep": "11",
      "nom_dep": "AUDE",
      "chef_lieu_dep": "11069"
    },
    {
      "dep": "66",
      "nom_dep": "PYRENEES-ORIENTALES",
      "chef_lieu_dep": "66136"
    }
  ]
}
```

Listing 2 – Exemple d'une ancienne région

```
{
  "_id": "occitanie",
  "chef_lieu_reg": "31555",
  "nom_reg": "Occitanie",
  "type": "region",
  "president": {
    "nom": "Delga",
    "prenom": "Carole"
  }
}
```

Listing 3 – Exemple de nouvelle région

3. Appropriation de la base de données

Vous répondrez aux questions suivantes en vous aidant de requêtes CURL :

1. lister les informations générales concernant le serveur couchdb, à l'aide du mécanisme GET
2. lister les informations générales concernant la base occitanie, à l'aide du mécanisme GET. Pouvez vous connaître le nombre de documents contenus dans la base occitanie ?
3. lister tous les documents de la BD
4. faire afficher le contenu d'un document. Quel est son numéro de révision ? Comment savoir si ce document a déjà été modifié ?

4. Définition de vues

Vous définirez des vues (en javascript) dans votre base pour consulter les documents décrivant les régions et les communes. Ce travail peut se faire soit en ligne de commande avec curl, soit avec l'interface Web Fauxton (depuis un navigateur à l'adresse `prodpeda-couchdb-2.infra.umontpellier.fr : 5984/_utils/`).

4.1 MAP seulement

1. donnez toutes les informations sur les régions (de type `old_region`) de la base
2. donner les noms (clés) et latitude et longitude de chaque commune
3. donner le code insee (clé), le département, la latitude et la longitude de MONTPELLIER (nom de la commune)
4. donnez le nom et le prénom de la présidente de la région Occitanie

4.2 MAP et REDUCE

1. donner le nombre de communes au total puis par département et enfin par région (`old_region`)
2. donner le nombre d'habitants par commune en 1985
3. donner le nombre d'habitants par département en 1985
4. donner le nombre d'habitants par région (anciennes régions) en 1985

4.3 Autres requêtes

1. donner les communes qui ont vu leurs populations décroître entre 1985 et 1995
2. donner les informations sur la nouvelle région Occitanie ainsi que sur les anciennes régions Languedoc-Roussillon et Midi-Pyrénées (forme de jointure)

5. Distribution de la base de données

Vous répondrez aux questions suivantes en vous aidant de requêtes CURL :

1. quels sont les nœuds mobilisés pour l'organisation du serveur, et la gestion de la base ?
2. Combien de partitions sont définies (avant recopie) ? Quel est le nombre de copies ? Combien de partitions répliquées sont définies ?
3. Comment savoir dans quelle(s) partition(s) se trouve un des documents de la base ?
4. Est ce que des copies de toutes les partitions sont présentes sur tous les nœuds ?
5. Est que toutes les lectures sont consistantes ?