

le cnam

Techniques de hachage

Exercices

CNAM Paris

nicolas.travers (at) cnam.fr

Cette section va permettre de mettre en pratique les tables de hachages à taille variable, afin de mieux appréhender l'élasticité d'une base de données distribuée sur le Cloud.

1.1 Hachage linéaire

Soit la liste de mots suivants à indexer :

parler, avoir, etre, aller, falloir, devoir, plus, moins, seul, lent, croissance, recession, vrai, france, inserer, pib, penser, humer

1.1.1 À chaque mot est attribué comme clé de hachage la valeur de sa première lettre

1.1.2 Créer une table de hachage linéaire de taille avec $p=0$ et $n=1$, avec 3 mots maximum par bloc de partition

1.2 Hachage Consistent - Anneau Chord

1.2.1 Création d'un réseau Chord

1.2.1 Construire le réseau Chord sur l'anneau de 16 nœuds (2^4) de la page suivante, dont la table de hachage distribuée contient dans chaque noeud 4 références (de 2^0 à 2^3 sauts).

1.2.2 Soit les 4 documents suivants contenant un ensemble de termes.

docId	terms
1	France, Récession, Ralentie, INSEE
2	USA, Finance, Paulson, Ralentie, Crise
3	Crise, Gouvernement, Ralentie, INSEE
4	PIB, France, Crise, Ralentie

Nous souhaitons distribuer pour chaque terme la liste des documents le contenant dans un réseau DHT (avec Consistent Hashing). Modéliser le réseau Chord de 16 noeuds avec la liste des docId pour chaque terme, avec la répartition suivante :

France : 1, Récession : 2, Ralentie : 3, INSEE : 4, USA : 5, Finance : 6, Paulson : 7, Crise : 8, Gouvernement : 9, PIB : 10

1.2.3 Le noeud 1 souhaite ajouter le document 5 contenant les mots suivants (Récession, PIB). Décrire le déroulement de l'insertion.

1.2.2 Interrogation du réseau

On souhaite trouver les documents à l'aide d'une requête avec deux mots¹ : *crise* \wedge *ralentie*

1.2.1 La requête est soumise au niveau du noeud 14. Décrire les étapes de l'évaluation de cette requête.

1.2.2 Donner le nombre de sauts nécessaire en tout, pour avoir la réponse à la requête.

1.2.3 Réseau Chord non complet

En pratique, le réseau n'est pas complet (nombre de serveur / nombre de noeuds maximum). Faire les mêmes exercices avec les serveurs placés aux indices suivants : 1, 3, 5, 8, 10

1. Uniquement les documents contenant ces deux mots font parti du résultat

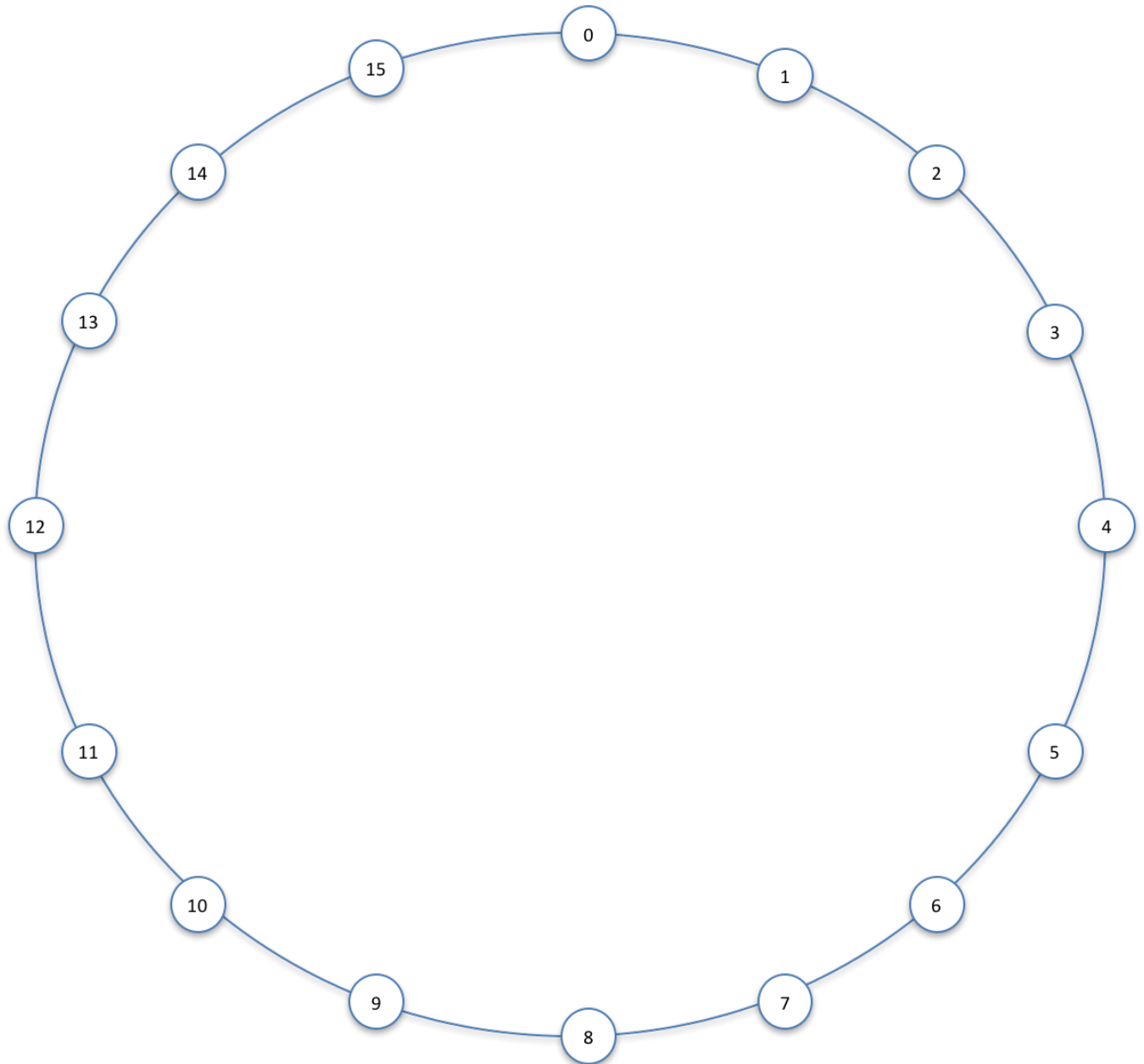


FIGURE 1.1 – Anneau de 16 nœuds