

## Exercices du cours ISI « Bases de Données Clouds 2011»

Witold Litwin

### Concepts Généraux

1. Expliquez le concept de « cloud ».
2. Quels sont les principaux « clouds » publiques à ce jour.
3. Que veut dire le terme « Data Center »
4. Regardez sur le Web une vidéo présentant un Data Center d'Azur
5. Justifiez les perspectives de cette technologie pour les BDs.
6. Commentez les définitions pratique et théorique d'une « BD Cloud »
7. En quoi consiste l'idée de Non SQL cloud DB ? Des exemples ?
8. Décrivez brièvement quelques clouds qui permettent de gérer BD Clouds en pratique.
9. Quoi et pourquoi de RamClouds, avec un commentaire spécifique pour la conjecture de J. Gray et le projet de J. Ousterhout & al en cours.
10. Comment les axiomes d'une SDDS évitent les points d'accumulation ?
11. En quoi diffère la partition SD par hachage de celle ordonnée ?

### LH\*

12. A quoi sert le niveau de la case dans LH\* ?
13. Comment LH\* évite les points d'accumulation ?
14. Répondez à toutes les interrogations dans les diapos du cours.
15. On crée une SDDS  $F$  de type LH\* (basique) où  $b = 4$  et  $N=1$  initialement. On insère la série  $S$  de clés qui suit,  $S$  étant pseudo-aléatoire sans doublons. Montrez l'évolution du fichier, après chaque insertion, les IAM selon l'un et l'autre algorithme du cours et l'évolution de l'image du client. Indiquez les clés qui ont subi un ou deux renvois.  
$$S = 10, 3, 1, 12, 15, 27, 32, 45, 43, 78, 59, 11, 7, 8, 33, 51, 77, 37, 48, 56, 131, 75, 47, 31, 14, 8, 2, 35, 42, 52, 62, 72, 57, 67...$$
16. Démontrez formellement que LH\* ne peut générer au max que deux messages de renvoie (« forwarding, hops.. .»)
17. Quelles sont actuellement les différences/similitudes entre les concepts de scan et celui de Map/Reduce ?
18. Que veut dire le concept de « Elastic Map/Reduce » par Amazon ?

19. Comment LH\* réalise la phase Map et celle de Reduce pour une requête d'interrogation, quand chaque réponse d'une case doit être retournée directement au client.
20. Montrez l'exécution d'un scan sur  $F$  de l'ex. 11, calculant le nombre d'articles dans chaque case du fichier. Le scan est lancé par client  $C$  avec l'étendu  $N'$  de l'image = 3. Le Reduce utilise le protocole déterministe, en répondant directement au client.
21. Le client  $C$  souhaite connaître la valeur moyenne de la clé primaire dans  $F$ . Il demande que le résultat final soit rendu par la case 0. Montrez l'exécution de ce scan.

**RP\***

22. Créez une SDDS  $T$  de type RP\* en y introduisant 50 mots successifs de l'article proposant cette SDDS. Chaque mot est supposé devenir la clé primaire  $C$ . Supposez cette création sous les trois variantes de RP\*.
23. Montrez l'exécution de la requête pseudo-SQL `Select * from T where C like R%`. Le résultat devra être sûr pour le client.