

Préparation à l'examen de *NoSQL*

Wéry Benoît

9 novembre 2017

Chapitre 1

Vrai ou Faux

1. *Dans une entreprise, les données vivent souvent plus longtemps que les logiciels.* **Vrai**
blabla
2. *Dans une entreprise, les logiciels vivent souvent plus longtemps que les données.* **Faux**
coucou
3. *Le passage du relationnel au NoSQL se fait généralement au profit d'une diminution des garanties relatives à la consistance des données.*
4. *Le passage du relationnel au NoSQL se fait généralement au profit d'une diminution de la quantité de données stockables.*
5. *Le passage du relationnel au NoSQL se fait généralement au profit de l'abandon de la possibilité de lire des données de manière concurrente.*
6. *Le passage du relationnel au NoSQL se fait généralement au profit d'une diminution des garanties relatives à la persistance des données.*
7. *Tout comme pour le relationnel, l'organisation des données en NoSQL suit un modèle mathématique rigoureux.*
8. *Le NoSQL est particulièrement adapté à des traitements de données de type OLTP*
9. *Le NoSQL est particulièrement adapté à des traitements de données de type OLAP.*

10. *Toutes les bases de données de type NoSQL satisfont les propriétés ACID.*
11. *Un système distribué peut toujours garantir la consistance des données sur tous ses nœuds.*
12. *Le mouvement de l'Open Data consiste à fournir librement des données récoltées pour permettre à la communauté de les analyser.*
13. *Le mouvement de l'Open Data à créer des logiciels open source permettant d'analyser des données massives (big data).*
14. *Google File System (GFS) est un moteur de base de données NoSQL.*
15. *Apache Hadoop est une implémentation propriétaire de MapReduce, commercialisée par Oracle.*
16. *L'intégration des données dans une seule base pour les partager entre plusieurs applications permet d'obtenir les meilleures garanties en terme de préservation de l'intégrité des données.*
17. *Limiter l'accès aux bases de données d'une entreprise à une seule application qui offre une API aux autres permet de rendre un changement de leur structure plus facile.*
18. *Une base de données NoSQL (clé-valeur, document et colonne) stocke des agrégats que l'on peut comparer aux tables du modèle relationnel*
19. *Une collection de paires clé-valeur peut être assimilée à une table relationnelle à deux colonnes dont la colonne représentant les clés est la clé primaire de la table.*
20. *Supprimer la valeur associée à une clé est l'une des trois opérations de base que l'on peut réaliser sur une base de données clé-valeur.*
21. *Dans une base de données clé-valeur, il est généralement prévu de rechercher toutes les clés dont les valeurs satisfont une certaine propriété.*

22. *Il est possible d'imposer des contraintes sur les domaines des valeurs des paires clé-valeur d'une base de données clé-valeur.*
23. *Distribuer les données sur un cluster de machines fait partie des éléments mis en place dans le monde NoSQL.*
24. *Il est possible de faire du sharding de données pour une base de données se trouvant sur une machine unique.*
25. *Le sharding permet de récupérer les données en cas de corruption grâce à un stockage redondant de ces dernières sur plusieurs serveurs pouvant être physiquement à des endroits différents.*
26. *Réplication de données et sharding sont incompatibles.*
27. *La réplication master-slave offre la propriété de résilience à la lecture.*
28. *En utilisant une réplication master-slave, les données deviennent complètement inaccessibles une fois que le master tombe.*
29. *La consistance des données est plus compliquées à garantir avec une réplication master-slave qu'avec une réplication peer-to-peer.*
30. *La consistance des données est plus compliquées à garantir avec une réplication peer-to-peer qu'avec une réplication master-slave.*
31. *En utilisant une réplication master-slave, une lecture sur le master assurera toujours d'obtenir les données les plus récentes*
32. *Les buckets de Riak permettent de segmenter les données en plusieurs collections d'agrégats.*
33. *On ne peut pas stocker des arbres binaires comme valeurs avec Redis.*

34. *Redis garantit la persistance de données.*
35. *Les bases de données orientée colonnes optimisent le stockage disque pour des tables qui contiennent de nombreuses lignes.*
36. *Une base de données orientée colonnes est très adaptée lorsqu'on a plus d'opérations d'écriture que de lecture.*
37. *Une base de données orientée colonnes est très adaptée lorsqu'on a plus d'opérations de lecture que d'écriture.*
38. *Une base de données orientée colonnes est un map à deux niveaux.*
39. *Dans une base de données orientée colonnes, les familles de colonnes sont de préférence définies une fois pour toute lors de la création de la table.*
40. *L'avantage de l'utilisation de colonnes plutôt que de lignes est d'offrir une vitesse d'écriture plus grande de nouveaux enregistrements.*
41. *L'avantage de l'utilisation de colonnes plutôt que de lignes est d'offrir un meilleur taux de compression des données stockées.*
42. *L'avantage de l'utilisation de colonnes plutôt que de lignes est d'offrir de meilleures performances lors de la lecture de tous les enregistrements d'une table.*
43. *Une base HBase peut servir d'input/output de MapReduce (Hadoop)*
44. *Une base HBase peut servir de fichiers avec GFS (Google File System)*
45. *Une base de données orientée graphe stocke deux collections d'agrégats appelés nœuds et arêtes.*

46. *La suppression d'un nœud dans une base de données orientée graphe implique la suppression de toutes les relations partant et arrivant sur ce nœud.*
47. *Il est impossible de stocker une liste de personnes dans une base de données orientée graphe*
48. *SPARQL est un langage de requêtes générique permettant d'interroger n'importe quelle base de données NoSQL.*
49. *Gremlin est un langage de requêtes générique permettant de décrire des traversées de graphe.*
50. *Neo4j supporte les transactions ACID.*
51. *Les bases de données orientée graphe sont très adaptée pour le sharding.*
52. *OrientDB offre la possibilité d'utiliser le sharding de données.*
53. *Une approche pessimiste de la consistance des données consiste à se limiter à un serveur unique pour le stockage des données.*
54. *Garantir la consistance de lecture empêchera tout conflit de type write-write.*
55. *Garantir la consistance de mise à jour empêchera tout conflit de type read-write.*
56. *Garantir la consistance de réplication est impossible avec un système peer-to-peer.*
57. *Si mes données sont répliquées sur quatre nœuds, avec $W = 2$, il suffit de lire deux nœuds pour lire l'information la plus à jour.*
58. *Si mes données sont répliquées sur quatre nœuds, il suffit d'impliquer $W = 2$ nœuds dans l'écriture pour assurer une consistance des données.*

59. *L'utilisation d'un timestamp comme version stamp est moins lourd à déployer que d'utiliser un GUID (Globally Unique Identifier).*
60. *L'utilisation d'un GUID (Globally Unique Identifier) comme version stamp est moins lourd à déployer que d'utiliser un timestamp.*
61. *L'utilisation d'un GUID (Globally Unique Identifier) comme version stamp permet de retrouver la version la plus récente d'une donnée.*
62. *Le write optimiste est très cher à implémenter dans un modèle clé-valeur.*
63. *Dans une base de données orientée colonnes, les données transitent par plusieurs espaces de stockage avant leur destination finale permanent.*
64. *Il est possible d'utiliser de la réplication master-slave avec une base de données orientée graphe pour rendre les lectures plus performantes.*
65. *Les bases de données orientée documents permettent d'effectuer des transactions atomiques au niveau d'un document.*
66. *Les bases de données orientée documents permettent d'effectuer des transactions atomiques au niveau d'une collection.*
67. *On peut changer le modèle d'une base de données NoSQL entre clés-valeurs, colonnes et documents, tout en garantissant exactement le même ensemble de propriétés.*
68. *Le passage vers le NoSQL permet de se passer des ORMs.*
69. *Le NoSQL est très adapté pour stocker des données très uniformes.*
70. *Le passage du relationnel au NoSQL rend les calculs à effectuer sur les données plus lents suite à un éventuel cout de transfert des données au sein du cluster.*

- 71. *L'opération Map s'applique à une collection de documents et renvoie une collection modifiée.*
- 72. *L'opération Reduce produit un résultat unique.*
- 73. *ElasticSearch est une base de données NoSQL.*
- 74. *ElasticSearch possède des caractéristiques similaires aux bases de données NoSQL.*
- 75. *Une migration de données en relationnel ou en NoSQL implique toujours un stockage physique de données redondantes.*