



Examen du 21/12/2018

Durée : 2h00

Aucun document n'est autorisé. L'usage de la calculatrice est interdit. Les questions faisant apparaître le symbole ♣ peuvent présenter zéro, une ou plusieurs bonnes réponses. Les autres ont une unique bonne réponse. Des points négatifs pourront être affectés à de très mauvaises réponses.

1 MongoDB

Question 1 MongoDB utilise le protocole de validation à 2 phases (2PC)

- ☒ A Vrai ☐ B Faux

Question 2 ♣ Le mode de fragmentation (sharding) de MongoDB est caractérisé comme

- ☐ A diagonal ☒ B horizontal ☐ C vertical et horizontal ☐ D vertical

Question 3 ♣ Avec MongoDB, on utilise le paramètre `--configsvr`

- ☒ A lors de la spécification du répertoire d'un serveur mongo
☐ B lors de la mise en place d'un cluster fragmenté avec réplication
☒ C lors de la mise en place d'un replica set
☐ D lors de la mise en place d'un cluster fragmenté sans réplication

Question 4 ♣ On utilise la commande `rs.initiate()`,

- ☐ A à chaque lancement d'une requête d'agrégation
☐ B à chaque démarrage d'un shard
☐ C à la création d'un index secondaire
☐ D lors de la création du répertoire contenant les données d'une base mongodb
☒ E Aucune de ces réponses n'est correcte.

Question 5 ♣ L'exécutable `mongos` gère

- ☒ A le routage des requêtes
☒ B la fusion des résultats intermédiaires d'une requête
☒ C la distribution des données
☐ D la synchronisation de la réplication

Question 6 A quoi correspond la notion de chunk lors de la fragmentation dans mongoDB

Question 7 Le système de gestion de base de données MongoDB ne supporte que des indexations sur le champ `_id`

- ☒ A Faux ☐ B Vrai

2 Distribution

Question 8 A quoi correspond (fournir le nom et sa signification) le symbole ✕



Question 9 Indiquer 3 pannes que l'on retrouve dans SGBD Distribué et pas un SGBD centralisé

Question 10 Dans le protocole de validation à 2 phases (2PC), à quoi correspond le problème du blocage

Question 11 ♣ Dans une base de données distribuée, la notion de transparence concerne

- ☐ A le partitionnement ☐ B la localisation de données sur le cluster
☒ C la réplication

Question 12 Dans l'optimisation de requêtes d'un SGBD distribué, citer 2 contraintes à prendre en compte que l'on ne rencontre pas dans un système centralisé

Question 13 Le protocole de validation à 2 phases (2PC) est indispensable dans le cas d'une transaction local d'un SGBD distribué

- ☐ A Faux ☒ B Vrai

Question 14 A quoi correspond un mode semi-synchrone de réplication

Question 15 ♣ Les opérateurs disponibles dans la phase 1 de 2PC sont:

- ☐ A commit T ☐ B prepare T ☒ C ready T ☐ D abort T ☒ E no T

Question 16 Fournir 3 approches différentes pour gérer la réplication dans un SGBD distribué

Question 17 L'architecture de calcul distribuée prédominante du cloud computing est

- ☐ A shared-disk ☒ B shared-nothing ☐ C shared-memory

Question 18 Fournir au moins 2 différences entre un SGBD parallèle et un SGBD distribué

Question 19 Le fait de découper une requête en plusieurs morceaux et de les exécuter en parallèle sur plusieurs serveurs correspond au parallélisme

- ☐ A micro-query ☒ B intra-query ☐ C inter-query

Question 20 Le fait d'exécuter en parallèle la même requête sur plusieurs serveurs correspond au parallélisme

- ☒ A inter-query ☐ B intra-query ☐ C micro-query

3 Optimisation de requêtes

Question 21 La commande ANALYSE lit l'ensemble des tuples de la table analysée pour construire ses statistiques.

- ☐ A Faux ☒ B Vrai

Question 22 On dispose d'une base sans index ayant le schéma suivant: `etudiant(numEtud, nom, prenom, adresse)` `examen(numEtud, codeCours, note)`. La requête ci-dessous s'exécute trop lentement. Que proposez-vous?

```
SELECT numEtud FROM etudiant WHERE numEtud IN (  
SELECT numEtud FROM examen WHERE note < 10 );
```



Question 23 En supposant les données uniformément réparties dans les tables, donnez une estimation de la taille la table intermédiaire dans ce calcul de jointures $(R \bowtie S) \bowtie T$.

(Vous pouvez donner une expression sans faire explicitement le calcul)

	R(a,b)	S(b,c)	T(c,d)
nombre de tuples	300	1000	1000
nombre de valeurs distinctes pour b	30	500	
nombre de valeurs distinctes pour c		100	200

Question 24 Le meilleur plan d'exécution est celui qui fait les selections le plus tôt possible.

☒ A Faux ☐ B Vrai

Question 25 Donnez un exemple d'opérateur qui ne permet pas de faire du pipeling avec les autres opérateurs du plan d'exécution.

Question 26 Expliquez une situation où la matérialisation d'un résultat intermédiaire peut améliorer les performances.

Question 27 L'ordre des jointures d'évaluation peut influencer le résultat d'une requête.

☐ A Vrai ☐ B Faux

Question 28 On dispose d'une base sans index ayant le schéma suivant: `soiree(idSoiree, lieu, prixEntree)` `participe(personne, idSoiree, deguiselement)`.

La requête ci-dessous s'exécute trop lentement. Que proposez-vous?

```
SELECT DISTINCT idSoiree FROM soiree S WHERE  
( SELECT count(*) FROM participe P WHERE deguiselement = 'citrouille' AND  
S.idSoiree = P.idSoiree ) >= 100;
```

Question 29 Voici une ligne du résultat d'un EXPLAIN. Expliquez à quoi correspondent les valeurs 219.74, 226.33 et 16.

Hash Anti Join (cost=219.74..226.33 rows=16 width=5)

Question 30 Donnez la complexité de l'algorithme de jointure nested loop calculant $R \bowtie T$ en fonction de n_T , n_R , B_T et B_R où n_X représente le nombre d'enregistrements de X et B_X le nombre de blocs utilisés pour stocker X .

Question 31 L'ordre des jointures d'évaluation peut influencer le temps de calcul du résultat d'une requête.

☐ A Faux ☐ B Vrai

Question 32 La commande EXPLAIN estime le nombre d'enregistrements renvoyés par une requête. Ce nombre est toujours supérieur ou égal au nombre réel d'enregistrements renvoyés.

☐ A Vrai ☐ B Faux

Question 33 Donnez un exemple de situation où l'algorithme de merge join est plus avantageux que celui de hash join.

Question 34 ♣ Etant donnée une table `R(a,b,c)`, quel(s) index peu(ven)t être utile(s) dans l'évaluation de la requête `select a from R where b>10; ?`

☐ A index hash sur a ☐ B index hash sur b ☐ C index btree sur b
☒ D index btree sur a ☐ E aucun

Question 35 Expliquez en quelques lignes le principe de l'algorithme de jointure "merge join".



La table $R(a,b,c)$ est munie d'un index btree sur l'attribut a et contient 20000 enregistrements dont un qui a la valeur 11 pour a , un qui a la valeur 10 pour b , et trente avec la valeur 10 pour a . Dans chacun des cas, quel opérateur est le plus adapté à l'évaluer la requête donnée?

Question 36 ♣ `select a,b from R where c=523;`

- ☐ A index only scan ☐ B bitmap index/heap scan ☐ C index scan
☐ D seq scan

Question 37 ♣ `select a from R where a<150;`

- ☒ A bitmap index/heap scan ☐ B index only scan ☐ C seq scan
☐ D index scan

Question 38 ♣ `select a,b from R where a=11;`

- ☐ A bitmap index/heap scan ☐ B index only scan ☒ C index scan
☐ D seq scan

Question 39 ♣ `select * from R where b=10;`

- ☐ A index only scan ☐ B seq scan ☐ C bitmap index/heap scan
☐ D index scan



Feuille de réponses :

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

← Codez votre numéro d'étudiant ci-contre de la manière suivante : un chiffre par ligne, le premier sur la première ligne, etc. Écrivez votre nom et prénom ci-dessous.

Nom et prénom :

Les réponses aux questions sont à donner exclusivement sur cette feuille : les réponses données sur les feuilles précédentes ne seront pas prises en compte.

QUESTION 1 : ☐ A ☐ B

QUESTION 2 : ☐ A ☐ B ☐ C ☐ D

QUESTION 3 : ☐ A ☐ B ☐ C ☐ D

QUESTION 4 : ☐ A ☐ B ☐ C ☐ D ☐ E

QUESTION 5 : ☐ A ☐ B ☐ C ☐ D

QUESTION 6 :

☐ A ☐ B

.....
.....

QUESTION 7 : ☐ A ☐ B

QUESTION 8 :

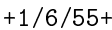
☐ A ☐ B ☐ C

.....
.....

QUESTION 9 :

☐ A ☐ B ☐ C

.....
.....
.....



A B C

A B C

A B C D

<div style="border-bottom: 1px dashed black; margin-bottom: 10px;"></div> <div style="border-bottom: 1px dashed black;"></div>
--

A

B

A B C

<div style="border-bottom: 1px dashed black; margin-bottom: 5px;"></div> <div style="border-bottom: 1px dashed black;"></div>

A

B

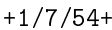
C

D

E

A B C

A B C



A B C

<div style="border-bottom: 1px dashed black; margin-bottom: 5px;"></div> <div style="border-bottom: 1px dashed black;"></div>

A

B

C

A B C

A B

A B C D E

[illegible]

A B C

A

B



QUESTION 25 :

☐ A ☐ B ☐ C

.....

.....

.....

QUESTION 26 :

☐ A ☐ B ☐ C

.....

.....

.....

.....

QUESTION 27 : ☐ A ☐ B



QUESTION 28 :

☐ A ☐ B ☐ C ☐ D ☐ E

.....

.....

.....

.....

.....

.....

.....

QUESTION 29 :

☐ A ☐ B ☐ C ☐ D

.....

.....

.....

.....

.....

.....



A B

<div style="border-bottom: 1px dashed black; margin-bottom: 5px;"></div> <div style="border-bottom: 1px dashed black; margin-bottom: 5px;"></div>

A B

A

B

A B C

[illegible]

A

B

C

D

E



QUESTION 35 :

☐ A ☐ B ☐ C

.....

.....

.....

.....

.....

.....

.....

QUESTION 36 : ☐ A ☐ B ☐ C ☐ D

QUESTION 37 : ☐ A ☐ B ☐ C ☐ D

QUESTION 38 : ☐ A ☐ B ☐ C ☐ D

QUESTION 39 : ☐ A ☐ B ☐ C ☐ D

