

Audit et optimisation MySQL 5

**Bonnes pratiques
pour l'administrateur**

Pascal Borghino

Olivier Dasini

Arnaud Gadal

© Groupe Eyrolles, 2010, ISBN : 978-2-212-12634-1

EYROLLES



Table des matières

CHAPITRE 1

Gérer une situation d'urgence avec MySQL..... 1

À chaque degré d'urgence sa panoplie d'outils 1

Temps de résolution : dix minutes 2

Étape 0 : informez et communiquez ! 2

Ne restez pas seul et discutez avec d'autres administrateurs 3

Consultez les informations système : journal d'erreurs, activités disques et processeur... 3

Tentez de vous connecter à la base 4

ATTENTION *Précautions à prendre avec une table MyISAM corrompue* 4

ASTUCE *Défilement page par page pour SHOW FULL PROCESSLIST* 5

À SAVOIR *Éviter l'empilement des requêtes et décrypter le SHOW PROCESSLIST* 5

REMARQUE *Se référer aux chapitres concernés* 6

Supprimer les requêtes les plus lourdes 6

PRATIQUE *Supprimer des requêtes rapidement* 7

Éviter que l'authentification des utilisateurs repose sur un DNS : l'erreur unauthenticated user 8

Consulter son système de surveillance 8

Tranche de vie d'une campagne marketing improvisée 9

Temps de résolution : une heure 10

La chasse aux requêtes lentes 10

À LIRE *Un chapitre consacré à l'étude des journaux* 10

Réécrire les requêtes trop coûteuses 10

ASTUCE *Les tables statiques à la rescousse* 11

Les problèmes de réplication 12

Temps de résolution : une journée 13

LIRE *Chapitre 8 consacré à la réplication* 13

PRATIQUE *Un problème peut en cacher un autre* 14

Conseils généraux face à l'urgence 15

Tirer profit du passé 15

Anticiper les problèmes 15

ATTENTION *Modifications à chaud, en production* 15

<i>L'entraînement à l'urgence</i>	16
<i>Enregistrer les données de l'incident</i>	16
<i>L'état d'esprit à adopter dans l'urgence</i>	16
<i>Trouver de l'aide</i>	17

CHAPITRE 2

Choisir son serveur MySQL 19

La mise à jour matérielle, une étape nécessaire ? 20

Les questions à se poser 20

ASTUCE *Identifier les goulets d'étranglements* 20

ASTUCE *Optimiser son serveur* 21

JARGON *Scaling up, scaling out et scaling back* 22

Du 64 bits oui... mais partout ! 22

Limites des systèmes 32 bits 23

ATTENTION *Ne soyez pas trop gourmands !* 23

RESSOURCES EN LIGNE *D'avantage de détails sur l'adressage mémoire* 24

Choisir ses processeurs 25

État des lieux 25

Les solutions face aux problèmes de montée en charge 25

JARGON *Architecture SMP vs NUMA* 25

REMARQUE *La compétition omniprésente entre les différents acteurs* 26

À LIRE ÉGALEMENT 27

Choisir son processeur : les critères de choix 27

Quelle est l'utilisation actuelle de vos processeurs ? 27

À SAVOIR *MySQL et la gestion des threads* 27

Fréquence vs nombre de cœurs 28

JARGON *OLTP, OLAP : deux catégories de systèmes à gérer différemment* 28

Benchmarks, encore et toujours 29

À LIRE ÉGALEMENT *Mesurer les performances de son système* 29

RAPPEL *Configurer son serveur MySQL* 30

ASTUCE *Pour aller plus loin* 31

Choisir ses disques et son système RAID 31

À SAVOIR *Temps d'accès mémoire vs temps d'accès disques* 32

Temps d'accès versus taux de transfert 32

À RETENIR *Lecture/écriture aléatoire ou séquentielle* 33

La technologie RAID 34

ATTENTION *La réplication et la montée en charge des écritures* 34

Les principaux niveaux de RAID 34

REMARQUE *Les opposants au RAID 5* 36

RESSOURCES EN LIGNE *D'autres niveaux de RAID existent* 37

Les deux implémentations du RAID : logicielle et matérielle 38

REMARQUE *Carte contrôleur RAID, force et faiblesse à la fois* 38

<i>Intérêt du cache sur une carte contrôleur RAID</i>	39
RAPPEL <i>Le cache de requêtes en amont de la carte RAID</i>	40
BON À SAVOIR <i>Les outils pour vérifier les réglages de sa carte contrôleur</i>	41
<i>Indispensable batterie</i>	42
EN PRATIQUE <i>Durée de vie de la batterie d'une carte RAID</i>	42
<i>Le cache interne des disques : une arme à double tranchant</i>	42
JARGON <i>innodb_flush_method = O_DIRECT</i>	43
À LIRE ÉGALEMENT	44
Les SSD : futur hit ?	44
MySQL et la mémoire	45
Comment MySQL utilise-t-il la mémoire ?	46

CHAPITRE 3

Les moteurs de stockage 49

Mécanismes d'un moteur de stockage	49
Installation et suppression d'un moteur	52
ATTENTION <i>Suppression d'un moteur de stockage utilisé par une table</i>	54
Les forces en présence : moteurs utilisés par l'application	54
B.A.-BA <i>Créer ses tables à partir de l'existant</i>	55
CONVENTIONS TERMINOLOGIQUES <i>Base de données, serveur, instance, schéma</i>	57
Les critères de choix d'un moteur	57
Moteurs disponibles : InnoDB, MyISAM, Merge, Memory, Archive	58
Le moteur InnoDB	58
B.A.-BA <i>Les propriétés ACID</i>	60
BON À SAVOIR <i>Le MVCC</i>	60
ASTUCE <i>Sortir une table d'un tablespace partagé</i>	63
POUR ALLER PLUS LOIN	63
MyISAM	63
<i>Mécanismes internes de MyISAM et formats de stockage</i>	67
B.A.-BA <i>Chaud, froid ou tiède ?</i>	67
RAPPEL <i>Mécanisme d'une commande ALTER TABLE</i>	68
Le moteur MERGE pour agréger plusieurs tables MyISAM	69
Le moteur MEMORY (anciennement HEAP)	71
Le moteur ARCHIVE pour un archivage compressé	72
Autres moteurs	72
<i>XtraDB</i>	72
<i>Falcon</i>	73
<i>Federated</i>	73
<i>Example</i>	73
<i>Blackhole</i>	73
<i>CSV</i>	74

<i>IBMDB2I</i>	74
<i>NDB (Network Database)</i>	74
Moteurs communautaires et autres	75
<i>Maria</i>	75
<i>PBXT</i>	75
<i>BLOB Streaming Engine (MyBS)</i>	76
<i>Mdbtools</i>	76
<i>Kickfire</i>	77
<i>TokuDB</i>	77
<i>Spider</i>	77
<i>Rethinkdb</i>	78

CHAPITRE 4

Surveiller son serveur MySQL..... 81

Où trouver les informations pertinentes ?	82
Variables système et variables de statut	82
DÉFINITION <i>Variables système ou de statut</i>	82
ALTERNATIVE <i>Récupérer les variables système ou de statut</i>	82
Quels outils choisir ?	83
À SAVOIR <i>Variables système et my.cnf</i>	83
ATTENTION <i>Une valeur peut en cacher une autre</i>	84
À SAVOIR <i>Différence entre un client et un outil MySQL</i>	84
Intérêt des outils de surveillance	84
B.A.-BA <i>key_buffer_size</i>	85
Outils et commandes fournis par MySQL	85
ATTENTION <i>Variables globales vs variables de session</i>	86
<i>Catégorie General</i>	87
B.A.-BA <i>MySQL vs mysqld</i>	88
ASTUCE <i>Les jokers dans les commandes MySQL : % et _</i>	89
LE SAVIEZ-VOUS <i>Deux descripteurs de fichiers pour une table MyISAM</i>	89
<i>Catégorie Performance</i>	91
À SAVOIR <i>Le cache de requête (Query Cache)</i>	91
RAPPEL <i>Le cache MyISAM</i>	93
RAPPEL	94
À SAVOIR <i>Ajuster la taille du cache d'index</i>	94
ATTENTION <i>Sortir un serveur client de la liste noire d'un serveur MySQL</i>	99
À SAVOIR <i>Droits nécessaires aux commandes SHOW STATUS et SHOW VARIABLES</i>	99
ASTUCE <i>Optimiser et analyser une requête avec USE INDEX/IGNORE INDEX</i>	104
La commande SHOW ENGINE INNODB STATUS	104
À SAVOIR <i>Différence entre mutex et sémaphores</i>	106
ASTUCE <i>Créer un deadlock délibéré</i>	108
À LIRE ÉGALEMENT <i>Le MVCC</i>	109

INFORMATION_SCHEMA	113
Connaître et savoir exploiter les outils de surveillance	114
Qu'est-ce que la performance ?	115
À LIRE <i>Technologie du disque</i>	116
LVM : la gestion des volumes logiques	117
JARGON <i>Transactionnel et cohérence, quelles différences ?</i>	118
B.A.-BA <i>Les différents types de sauvegardes (backups)</i>	119
Étude de cas : analyse d'un serveur MySQL	120
RAPPEL <i>write-through/write-back</i>	123
Mesurer l'activité du serveur	124
Les outils système	124
<i>La commande iostat</i>	125
<i>La commande vmstat</i>	126
<i>Les commandes netstat et mpstat</i>	127
ALTERNATIVE <i>oprofile, dtrace, ftrace et filefrag</i>	127
Outils d'audit : MySQLTuner et mysqlreport	127
ASTUCE <i>Surveiller son serveur à distance</i>	130
Outils d'analyse temps réel : mytop, mtop, innotop et maatkit	131
Évaluer les performances d'un système	131
JARGON <i>Le smoke test, un test aux limites</i>	132
MÉTHODE <i>Dimensionnement : les bons tests</i>	135
Bien dimensionner un système (capacity planning)	136
À SAVOIR <i>La notion de seuil</i>	138
À LIRE <i>La montée en charge matérielle</i>	138
À LIRE <i>Pour aller plus loin dans le domaine</i>	139

CHAPITRE 5

Exploiter les journaux de MySQL **141**

Le journal des erreurs	142
ASTUCE <i>Rotation des journaux avec logrotate</i>	143
Identifier et résoudre les problèmes	143
Modifier le tablespace ou les journaux d'InnoDB	143
Paramètre incorrect dans le fichier de configuration	144
Erreurs liées à la réplication	146
Erreurs diverses	147
ATTENTION <i>Effets de bord de l'option myisam_recover</i>	148
Le journal des requêtes lentes	148
Principe de fonctionnement	148
ATTENTION <i>Effets de bord de l'option log_queries_not_using_indexes</i>	149
ALTERNATIVE <i>Autres outils d'analyse</i>	151
Journaliser dans une table	152

Le journal général des connexions et requêtes	153
Exemples d'utilisations de la journalisation générale ?	154
La journalisation binaire	155
REMARQUE <i>L'option <code>sync_binlog</code> peut avoir un impact sur les performances</i>	155
TRANCHE DE VIE <i>La technique de Point In Time Recovery en pratique</i>	157
REMARQUE <i>Taille du journal binaire en fonction du mode de journalisation</i>	158
Bonnes pratiques	160

CHAPITRE 6

Optimiser sa base de données : du schéma aux requêtes 163

Conception de la base de données	163
Normalisation/dénormalisation	164
BON À SAVOIR <i>La normalisation</i>	164
OUTILS <i>Logiciels de modélisation</i>	165
<i>Ajouter des colonnes dans une table</i>	165
<i>Création de tables d'agrégation</i>	166
<i>Création de schémas orientés</i>	167
Des types de données ajustés	167
MÉTHODE <i>Un type optimal à un moment donné</i>	168
Les jointures	169
Les index	172
Index B-tree	173
B.A.-BA	173
BON À SAVOIR <i>Index Fulltext (Plaintext)</i>	174
Index B+tree	175
BON À SAVOIR <i>La table de hachage</i>	177
Index hash	177
ALTERNATIVE	180
Optimisation des requêtes	181
Connaître l'optimiseur pour mieux le comprendre	181
B.A.-BA <i>La sélectivité</i>	181
BON À SAVOIR <i>Optimiseurs à base de règles ou de coût (cost based ou rules based)</i>	182
La commande EXPLAIN pour analyser l'exécution des requêtes	182
PRATIQUE <i>Visualiser le plan d'exécution d'un DELETE ou d'un UPDATE</i>	183
REMARQUE <i>Attention à la fonction RAND()</i>	185
OUTILS <i>Représentation graphique du plan d'exécution avec Maatkit</i>	186
BON À SAVOIR <i>Optimisation des index et réorganisation des tables</i> <i>avec ANALYSE TABLE et OPTIMIZE TABLE</i>	187
Exemple d'optimisation d'un plan d'exécution	187
<i>Indexer les premiers caractères d'une colonne</i>	189
<i>Index couvrant (covering index)</i>	190

<i>Préfixe d'index (leftmost prefix indexes)</i>	190
BON À SAVOIR <i>Pas de préfixes d'index pour les index hash</i>	190
<i>Taille des index</i>	190
Récapitulatif des bonnes pratiques d'optimisation des requêtes	190
<i>Découper les requêtes complexes en plusieurs plus simples</i>	191

CHAPITRE 7

Optimiser son serveur MySQL **193**

Tuning serveur : variables de session, variables globales, handlers 193

VOCABULAIRE *Cache et buffer (tampon)* 194

Les variables de session 195

read_buffer_size 195

read_rnd_buffer_size 195

sort_buffer_size 195

join_buffer_size 196

tmp_table_size et *max_heap_table_size* 196

Les variables globales au serveur 196

Le cache de table 196

Le cache de thread 197

Table_locks_immediate et *Table_locks_waited* 198

Aborted_clients 198

Aborted_connects 198

Les handlers 199

Exemple d'optimisation d'une requête 199

Les droits des utilisateurs 203

Optimisations pour InnoDB, MyISAM et MEMORY 204

Optimisation InnoDB 204

Optimisation MyISAM 205

Cache d'index multiples 206

Optimisation Memory 207

ATTENTION *Limiter la taille des tables* 207

RAPPEL *Limitations du moteur Memory* 207

Le cache de requêtes (query cache) 207

Gestion du cache de requêtes 210

ATTENTION *Taille du cache de requêtes* 210

Le partitionnement 211

Le partitionnement par RANGE 213

Le partitionnement par LIST 214

Le partitionnement par HASH 214

Le partitionnement par KEY 214

Partitionner sur différents disques 215

Partitionner sur différents disques avec MyISAM	215
BON À SAVOIR <i>Évolution du partitionnement</i>	215
CHAPITRE 8	
La réplication MySQL	217
Introduction à la réplication	217
Intérêt de la réplication	219
Le dimensionnement horizontal (scale out)	219
La sauvegarde à chaud (hot backup)	219
Le basculement automatique (Failover)	220
Redondance géographique	220
Le cas du décisionnel	220
Tester une nouvelle version de MySQL	220
À l'intérieur de la réplication	221
Mise en place de la réplication	221
Configuration du maître	222
ATTENTION <i>Mot de passe en clair</i>	222
BON À SAVOIR <i>Filtrage des données répliquées</i>	224
Configuration de l'esclave	225
REMARQUE <i>Mettre à jour l'esclave avant le maître</i>	226
ATTENTION <i>Ancienne méthode de configuration</i>	226
Configuration avancée de l'esclave	226
Commandes de la réplication	227
Sur l'esclave	228
ATTENTION <i>Conséquences d'un RESET SLAVE</i>	228
ASTUCE <i>Comment savoir si le serveur esclave a du retard ?</i>	230
<i>La commande SHOW SLAVE STATUS.</i>	231
Sur le maître	232
ASTUCE <i>Déconnexion d'un serveur esclave</i>	232
ATTENTION <i>RESET MASTER peut casser la réplication</i>	233
Problèmes liés à la réplication	233
IO_THREAD stoppé	234
SQL_THREAD stoppé	234
DANGER <i>Ignorer les erreurs peut provoquer des incohérences</i>	235
BON À SAVOIR <i>Tables temporaires et réplication</i>	236
Architectures de réplication avancées	237
REMARQUE <i>MySQL Cluster</i>	237
Dual master en actif/passif	237
DANGER <i>Une réplique n'est pas une sauvegarde</i>	238
<i>Configuration</i>	239
OUTILS <i>Supervision des serveurs</i>	239

<i>Exemple : switchover pour une mise à jour online des serveurs MySQL</i>	239
<i>Récapitulatif</i>	241
ALTERNATIVE <i>Commencer par le maître</i>	241
Dual master en actif/actif	241
<i>Récapitulatif</i>	244
Réplication circulaire (nombre de répliquions > 2)	244
Esclave relais	244
<i>Configuration</i>	246
<i>Récapitulatif</i>	247
Partitionnement adapté au décisionnel	248
<i>Configuration</i>	249
<i>Récapitulatif</i>	250
Bonnes pratiques	250
À SAVOIR <i>Le sharding</i>	252

CHAPITRE 9

Où trouver de l'aide ? 253

Trouver de l'aide en urgence	254
Les ressources internes	254
Les ressources externes	254
Les moteurs de recherche	254
Le support officiel MySQL	254
Les organismes externes	255
Trouver de l'aide hors contexte d'urgence	256
Formations	256
Où poser votre question ?	256
L'association LeMug	256
Les blogs	257
Les forums et mailing-lists MySQL officiels	257
Aller plus loin et enrichir ses connaissances	257
La blogosphère de la communauté	258
Les séminaires web	259
Outils et sources de MySQL	259
La conférence MySQL	259
Les certifications	260

Index..... 261

Avant-propos

Un ouvrage en français pour aller plus loin

L'idée de départ de cet ouvrage fut d'écrire ce que nous-mêmes aurions aimé trouver en librairie au rayon MySQL.

Le contenu que nous vous proposons aujourd'hui, et qui sera détaillé dans quelques paragraphes, est le fruit de la diversité de nos expériences respectives (un administrateur base de données, un expert MySQL et un architecte bases de données). En parcourant ce livre, vous profiterez de nos expériences acquises auprès de grands comptes de l'univers Internet, tels que Orange Business Services, Virgin Mobile ou encore Yahoo!

Nous souhaitons écrire un livre permettant de mieux comprendre à la fois MySQL en tant que pièce logicielle, mais aussi son lien avec le serveur physique sur lequel cette base de données est installée. Cette subtile interaction entre logiciel et matériel où interviennent plusieurs centaines de variables et autres paramètres, est susceptible d'atteindre de hautes performances lorsque tous agissent de concert.

Cependant, il ne s'agit pas uniquement de décrire les mécanismes responsables des multiples comportements de MySQL observés ; il faut surtout les expliquer.

Directement issu de nos vies professionnelles, cet ouvrage est un reflet condensé de nos différentes expériences. Ce que vous lirez ici, nous l'avons vraiment vécu. Les fortes charges, les pics d'affluence, les corruptions de données, les requêtes qui n'en finissent pas, une volumétrie qui explose, des répliquions en échec, des serveurs MySQL à l'agonie... nous y avons tous goûté !

Fort heureusement, les problèmes que nous venons d'évoquer ne sont pas une fatalité et les bonnes pratiques exposées dans cet ouvrage vous permettront de les éviter pour la plupart ou, en tout cas, de les gérer au mieux. Rassurez-vous, les différents chapitres qui vont suivre ne sont pas une suite de récits de cataclysmes MySQL ! Au con-

traire, il s'agit plutôt de prévenir ces écueils en adoptant les bonnes techniques qui vous permettront, à vous qui gérez ou utilisez au quotidien des bases MySQL, de les utiliser le mieux possible.

Le titre, *MySQL 5 - Audit et optimisation*, résume le cœur de l'ouvrage. En analysant techniquement un serveur MySQL lors de la phase d'audit, nous établissons un diagnostic qui permettra les optimisations appliquées sur cette machine. C'est ainsi que tous les chapitres apportent leur valeur ajoutée à cette thématique.

Nous ne connaissons pas d'ouvrage en français rassemblant l'étendue des thèmes évoqués ici. Le livre que vous tenez entre les mains n'est pas simplement du contenu pointu « en français ». Sa plus-value réside également dans notre fort intérêt pour MySQL. Nous sommes trois passionnés de cette base de données (blogueurs, association LeMug.fr) et notre but tout au long du livre est de vous faire partager cette passion tout en décortiquant les mécanismes de MySQL de façon claire et pratique. Encarts, schémas et exemples étayeront notre propos et vous guideront tout au long du chemin que nous vous avons tracé vers l'optimisation de vos serveurs MySQL.

À qui s'adresse cet ouvrage ?

Cet ouvrage est particulièrement destiné à tous ceux qui ont déjà des connaissances de base en MySQL et qui sont soumis à des problèmes liés à une activité croissante sur leurs bases de données.

Qu'il s'agisse d'une audience croissante, de données de plus en plus lourdes ou encore d'engranger rapidement de nouvelles connaissances pointues sans avoir à parcourir toute la blogosphère anglophone, ce livre saura vous apporter des solutions concrètes ou vous aiguiller vers la bonne façon de faire.

De par son contenu, cet ouvrage s'adresse plutôt aux administrateurs de bases de données, qu'ils soient spécialistes de MySQL ou pas, qu'aux développeurs. Cependant, les plus curieux d'entre eux trouveront des informations sur l'optimisation de leurs requêtes et sur les différences entre les moteurs de stockage, par exemple.

Pour les administrateurs MySQL, nous avons essayé d'apporter le maximum d'informations utiles à leur métier, notamment sur les outils que nous utilisons pour auditer un serveur, l'optimiser, mesurer ses performances, anticiper ses problèmes... Vous y trouverez donc tous nos meilleurs conseils.

Structure de l'ouvrage

L'ouvrage est découpé en huit chapitres qu'il est possible d'aborder indépendamment les uns des autres. Si certains thèmes sont abordés dans plusieurs chapitres, nous le signalons par des renvois appropriés. De notre point de vue, seul le chapitre « Gérer une situation d'urgence avec MySQL » devait apparaître en premier, l'ordre de lecture des autres chapitres n'ayant que peu d'importance.

Si vous êtes particulièrement intéressé par la réplication, vous avez notre bénédiction pour attaquer directement par le chapitre 7.

Outre les liens entre chapitres que nous venons d'évoquer et la possibilité de ne lire que certains chapitres, c'est en lisant l'intégralité de l'ouvrage que vous aurez une vue d'ensemble de toutes les techniques décrites, ce qui vous rendra plus efficace.

De plus, nous n'avons pas multiplié le nombre de chapitres artificiellement mais choisi au contraire de restreindre notre champs d'étude aux notions les plus importantes selon nous concernant l'audit et l'optimisation d'un serveur MySQL.

Nous débuterons donc par MySQL et l'urgence (**chapitre 1**). Que faire quand rien ne va plus ? Les bases de données sont un domaine où ne rien faire est parfois la meilleure des solutions... dans l'attente de trouver une réponse adaptée ; et où « tenter » quelque chose sans en analyser les conséquences est potentiellement dramatique et à éviter. Dix minutes, une heure ou une journée, voici les trois scénarios « d'urgence » que nous vous proposons de résoudre dans ce premier chapitre, le tout agrémenté de nos meilleurs conseils dans un tel contexte.

Le **chapitre 2**, « Choisir son serveur MySQL », met l'accent sur les impacts du matériel sur MySQL. La mémoire, les disques (RAID et cartes contrôleur, SSD), les processeurs (quantité, combien de cœurs ?), ces éléments ont une importance cruciale pour la base de données. Apprenez à les choisir.

InnoDB, MyISAM, Memory, Archive... une des spécificités de MySQL est de pouvoir connecter au serveur plusieurs « moteurs » de stockage, encore faut-il le faire à bon escient. Rendez-vous au **chapitre 3** pour mesurer quels sont les impacts du choix d'un moteur par rapport à un autre. Saviez-vous par exemple que le plan d'exécution de vos requêtes varie en fonction du moteur ? Comment fonctionnent les index ou les caches selon les moteurs, quelles sont les forces et les faiblesses de chacun ? Toutes les informations dont vous avez besoin pour choisir le moteur qu'il vous faut se trouvent ici.

Le **chapitre 4** vous aidera à connaître l'état de santé de votre serveur MySQL. Prendre le pouls d'un serveur nécessite de comprendre le fonctionnement des variables système et des variables de statut. Quels sont les liens qui les unissent et les outils qui les exploitent ? Vérifier l'état de santé du moteur InnoDB lui-même, pré-

voir la capacité maximale de votre système (*capacity planning*) et une étude de cas sont au programme.

Les journaux MySQL sont détaillés au **chapitre 5**. Ne sous-estimez pas la puissance des logs. Bien exploités, ils seront un allié de premier ordre non seulement en cas de coup dur mais également au quotidien. Jugez plutôt : journalisation des erreurs, logs des requêtes lentes, logs binaires pour la réplication, les logs sont partout.

Les **chapitres 6 et 7** baignent dans l'optimisation. Les index, les requêtes, l'étude du serveur lui-même (variables clés du fichier `my.cnf`), les caches en tous genres, mais aussi les optimisations propres à certains moteurs de stockage ainsi que le partitionnement sont autant de notions abordées dans ce chapitre.

La réplication se dévoile au **chapitre 8**. Cette fonctionnalité cruciale offerte par MySQL y est traitée avec beaucoup de détails. À quoi sert-elle, comment la mettre en place, quelles sont les commandes clés sur le maître et l'esclave ? Les différentes architectures disponibles et la résolution des problèmes liés à cette technologie sont également au programme.

Enfin, si d'aventure notre ouvrage vous laissait avec des questions en suspens, le **chapitre 8** indique où trouver de l'aide. Souhaitez-vous obtenir une aide de toute urgence ou simplement poser une question générale sur un forum ou une mailing-list adéquate ? Peut-être cherchez-vous les meilleurs blogs francophones ou anglophones pour enrichir vos connaissances ? Vous y trouverez nos meilleures adresses.

Remerciements

Plus attendue que la gloire éternelle liée à la rédaction d'un ouvrage technique, l'écriture de ces quelques paragraphes de remerciements fait sans doute partie d'une des motivations principales d'un auteur, nous nous plions donc volontiers à cette tradition.

Écrire ce livre à plusieurs mains a été pour moi une expérience enrichissante en grande partie grâce au professionnalisme de mes co-auteurs. Je tenais donc premièrement à remercier Arnaud et Olivier pour leur travail et leur bonne humeur. Ensuite, nos éditrices Muriel Shan Sei Fan et Sophie Hincelin, ainsi que Pascale Sztajn bok et Gaël Thomas de l'équipe éditoriale d'Eyrolles qui ont pu gérer avec patience nos légers retards et nous ont permis de nous focaliser sur le contenu et le fond. Ce livre n'aurait pas vu le jour non plus sans Véronique Loquet d'ALX communication, dont la passion et la connaissance du monde de l'open source ont su nous ouvrir les portes nécessaires. Encore merci Véronique !

Je voulais aussi remercier mes managers qui ont permis mon implication dans le monde de MySQL, dans l'ordre : Rémy, Bruno, Dana et Rusty... Je n'oublie pas non plus mes formidables amis de la communauté MySQL elle-même : merci Jeremy, Éric et toute la bande.

Enfin, merci à mes parents Daniel et Danielle qui seront toujours à la source de tous mes accomplissements...

Pascal Borghino

Je remercie mes deux amis, co-auteurs et gourous MySQL, Arnaud et Pascal, sans qui ce livre n'aurait jamais vu le jour. Je tiens également à y associer toute l'équipe de MySQL France, notamment Stéphane Varoqui et Serge Frezefond pour leur savoir infini. Merci à tous les passionnés du logiciel libre rencontrés au fil des années, sur les forums, les salons, dans les associations (LeMUG.fr, l'AFUP, l'April...) et sur mon blog (<http://dasini.net/blog/>).

Une tendre pensée pour mes parents Jean et Jocelyne, mes sœurs Karen et Linda, mon frère Floriant et à Kapinou (pour ta patience).

Enfin, je dédie ce livre à la mémoire de Valérie et Paul BERCHEL, ainsi que celle de Raymond DASINI.

Olivier Dasini

Je ne crois pas au destin mais plutôt au timing et un peu au hasard. Début 2006, je quittais Paris pour trois ans au détour d'un job d'ingénieur développement PHP pour le soleil de Sophia-Antipolis. Au bout d'un an, alors que je renforçais sérieusement mes connaissances MySQL, deux collègues là-bas ont probablement fait basculer ma carrière. Tout d'abord Gilles Oliveri (merci), qui a remarqué que je serais peut-être mieux employé à faire du MySQL que du PHP chez Orange, puis Cyril Scetbon (merci) qui m'a fait confiance et m'a accueilli dans la cellule Bases de données où j'ai pu apprendre, parfaire mes connaissances et surtout les mettre en pratique.

En 2008, ma passion pour MySQL m'a poussé à me rendre à la conférence annuelle de MySQL en Californie. Là-bas, Damien Seguy (merci) m'a présenté à Pascal (merci) qui m'a proposé de blogger pour dbnewz.com lors d'un MySQL Quizz Show alors que je lui reprochais de ne pas publier assez souvent ! Pas rancunier l'animal...

Quant à Olivier (merci), j'ai rencontré sa bonne humeur légendaire lors d'un forum PHP/MySQL il y a quelques années. À croire que toutes les routes mènent à MySQL...

Mi-2009, Pascal et Olivier m'ont proposé d'être le troisième homme pour la rédaction de cet ouvrage. J'ai accepté avant de savoir que j'aurais à les relire... Une pure folie... que je ne regrette pas ! Merci encore à tous les deux.

À toute l'équipe Eyrolles qui a pris soin de nous pendant toute cette aventure, merci. Enfin, à vous qui lisez ces quelques lignes, merci. Autant de motivation pour lire un avant-propos est très bon signe pour la suite.

Bonne lecture !

Arnaud Gadal

1

Gérer une situation d'urgence avec MySQL

Problème majeur ou incident moins dramatique, l'un comme l'autre méritent la plus haute attention. En effet, tous deux sont susceptibles d'être résolus rapidement ou au contraire d'impacter gravement la production. Voici nos conseils pour minimiser les conséquences de ces aléas qui ponctuent la vie d'une base de données.

Il serait illusoire de recenser ici l'ensemble des problèmes et des solutions associées que vous pourriez rencontrer sur un serveur MySQL. En revanche, il est possible de lister quelques bonnes pratiques en fonction du temps dont vous disposez face à un incident en production. Inutile en effet de se lancer dans un audit poussé du système si vous n'avez que quelques minutes pour rétablir une situation compromise. Cependant, avec un peu plus de temps, une heure voire une journée, le mode opératoire diffère complètement. Il est alors possible de rechercher plus longuement les causes du problème et d'activer si possible d'autres ressources en interne ou en externe, par exemple via du support. Bref, analyser le problème afin qu'il ne se reproduise plus.

À chaque degré d'urgence sa panoplie d'outils

Dix minutes, une heure et une journée sont les trois différents degrés d'urgence (arbitraires) que nous vous proposons dans ce chapitre. Cette distinction peut bien sûr être discutée, l'idée étant néanmoins de classer les problèmes par ordre de priorité. À bien y réfléchir, ces ordres de grandeur ne sont finalement pas si éloignés de la réalité.

Les incidents gravissimes (serveur MySQL qui ne redémarre pas, connexion à la base impossible sur un serveur en fonctionnement...) nécessitent une réponse immédiate de votre part car ils ont une incidence directe sur la production. Un délai de résolution de 10 minutes est souhaitable mais laissera malgré tout des traces dans vos tableaux de statistiques mesurant les temps de disponibilité de vos applications.

Légèrement moins urgents, les problèmes que l'on souhaite voir régler dans l'heure ne sont pas à prendre à la légère non plus. En effet, il peut s'agir de difficultés de réplication, de droits absents ou incomplets pénalisant des utilisateurs ou des scripts, de crash de tables, etc.

Enfin, les problèmes de performance sont susceptibles d'appartenir à la catégorie d'incidents dont la résolution, ou tout du moins le diagnostic, ainsi qu'un éventail de solutions adaptées, relève de la journée.

Temps de résolution : dix minutes

Bienvenue à ceux qui savent gérer leur temps ! Joueurs d'échecs, un plus. Voici une annonce qui sortirait sûrement du lot pour une entreprise souhaitant attirer dans ses filets un administrateur de bases de données (DBA). En effet, tel un joueur d'échecs confronté à une partie ultrarapide (5 minutes), le DBA en situation de crise dispose de très peu de temps pour tenter de renverser la situation. En cela, dix minutes constituent à la fois un laps de temps très court à l'échelle d'une journée, par exemple, mais on peut également considérer que c'est amplement suffisant pour détecter l'origine du problème. Nous parlons ici de détection, pas encore de résolution.

Étape 0 : informez et communiquez !

Identifiez la ou les personnes auxquelles vous devez référer en cas de problème. Celle-ci relayera le message si besoin. L'informatique est plus qu'une question de technique : un accident grave sur les bases passe rarement inaperçu et vous oblige à communiquer. Un incident, en particulier s'il est majeur, sera mieux vécu si les utilisateurs concernés par cet incident sont mis au courant plutôt que s'ils sont obligés de contacter eux-mêmes les services techniques pour tenter de comprendre pourquoi leur application ne répond plus. Prenez les devants et annoncez les conséquences : une réplication hors service implique au mieux un état figé des données sur certaines applications et au pire un arrêt de certaines d'entre elles, tout dépend de la robustesse du code sous-jacent.

De plus, indiquer aux utilisateurs qu'il existe un problème leur permet d'attendre votre feu vert avant de renouveler leurs opérations plutôt que de tenter des rafraîchissements ou validations supplémentaires qui n'arrangent rien.

Si vous ne pouvez d'ores et déjà annoncer un délai de rétablissement, dites-le. Outre le fait que les utilisateurs soient avertis au plus tôt, communiquer en amont vous permet également de rester concentré en réduisant le nombre de personnes cherchant à se renseigner auprès de vous (téléphone, messagerie interne, de vive voix...).

Ne restez pas seul et discutez avec d'autres administrateurs

En cas d'extrême urgence, partager ses observations, avec un administrateur système par exemple, est une bonne chose. Si le DBA pense en termes de bases (réplication, droits, tables corrompues), l'administrateur système a souvent une vue complémentaire (place disque, réseau, connectivité aux disques, scripts programmés) ; c'est souvent lui qui a installé, ou tout du moins participé à l'installation, de la machine sur laquelle le serveur MySQL s'exécute. Ses connaissances purement techniques ajoutées à la visibilité qu'il possède sur les machines utilisées par votre département en font un allié de choix en cas de crise.

Consultez les informations système : journal d'erreurs, activités disques et processeur...

Lorsqu'une alerte critique liée à la base de données est levée, une des premières actions à mener, si l'on ne dispose que d'informations vagues du type « système indisponible », est de consulter le journal d'erreurs de MySQL... À condition de pouvoir le faire bien sûr. Si l'un des disques système n'est plus accessible, l'incident bascule tout d'un coup du côté des administrateurs système. Vous êtes seul et occupez les deux fonctions ? Échangez votre casquette de DBA contre celle de l'administrateur système le temps que le problème matériel et/ou système soit résolu.

Les journaux d'erreurs sont un passage absolument obligé lorsqu'un incident relatif à la base de données survient. Sa simple lecture peut faire gagner un temps précieux ! Problème de réseau ou de réplication, arrêt du serveur, table corrompue, syntaxe incorrecte du fichier de configuration, impossibilité de démarrer un moteur de stockage, etc., l'éventail couvert par le fichier d'erreurs est vaste et bien souvent explicite.

Si sa lecture ne vous a pas suffi, ou s'il n'y a aucune information pertinente à l'intérieur, ce qui arrive, quelques commandes système basiques sont à exécuter :

- Tout d'abord `df -kh` pour s'assurer qu'aucune de vos partitions n'est pleine à 100 % ; cela dit le log d'erreurs vous l'aurait signalé.
- `iostat -dx 5` pour obtenir une vue sur l'activité des disques.
- `htop` ou `top` permettent quant à eux de repérer si vos processeurs, ou cœurs pour être plus précis, sont mobilisés et dans quelle proportion.

Espace disque disponible, activité processeur et activité du système disque sont quelques investigations qui doivent normalement contribuer à vous éclairer.

ATTENTION Précautions à prendre avec une table MyISAM corrompue

Le maître mot est ici : sauvegardez ! En effet, la documentation le stipule elle-même : il est possible que les procédures de réparation entraînent une perte de données :

► <http://dev.mysql.com/doc/refman/5.1/en/repair-table.html>

Si possible, veuillez donc sauvegarder les tables MyISAM concernées. En cas d'impossibilité de passer par un outil tel que *mysqldump* ou *mysqlhotcopy*, copiez directement les fichiers *.MYD*, *.MYI*, et *.frm* de la table concernée si le serveur est arrêté. Si ce dernier est en marche, la commande `FLUSH TABLES WITH READ LOCK` vous permet d'effectuer l'opération mais cela peut être coûteux en termes de performance.

Le plus important ici est de récupérer le fichier qui contient la base *.MYD*. En effet, si vous connaissez la définition de la table concernée, il sera possible de recréer un *.frm* lui correspondant. Le fichier *.MYI*, quant à lui, peut également être reconstruit : `REPAIR TABLE` utilisé avec l'option `USE_FRM` repartira justement du *.frm* et les index seront totalement créés à nouveau.

Pensez également aux logs binaires. Cette remarque est aussi valable pour un moteur tel que InnoDB. Si la partition sur laquelle se trouvaient vos tables MyISAM est perdue, récupérez vos sauvegardes et les logs binaires, puis réinjectez le tout dans MySQL et vous en serez quitte pour une bonne frayeur.

Attention toutefois à exclure de vos logs binaires l'éventuelle instruction SQL dévastatrice responsable du carnage... À ce sujet, reportez-vous au chapitre 8, consacré à la réplication.

Tentez de vous connecter à la base

Au-delà des outils système, pouvez-vous vous connecter à la base ? Pouvoir effectuer la commande `SHOW FULL PROCESSLIST` est très utile dans ce type de situation. Lorsque la base est chargée, il est efficace de passer par une ligne de commandes pour écrire dans un fichier texte le contenu de cette commande, exemple :

```
shell> mysql -uuser -ppassword -e "SHOW FULL PROCESSLIST;" > /tmp/sfp.txt
```

Il sera alors possible d'éditer facilement l'ensemble des requêtes SQL et de les conserver pour étude, le cas échéant, une fois la tempête passée.

L'utilisation de la commande `SHOW FULL PROCESSLIST` n'est pas forcément chose facile dans tous les cas : encore faut-il pouvoir se connecter !

À SAVOIR Éviter l'empilement des requêtes et décrypter le SHOW PROCESSLIST

Chaque serveur MySQL possède une limite maximale de connexions simultanées. Celle-ci est définie par la variable système `max_connections`. Certains scripts ou applications ne ferment pas correctement leur connexion à MySQL une fois le résultat de celle-ci exploité ; il en résulte un empilement des connexions. Celles-ci arborent alors le statut `sleep` mais occupent malgré tout un emplacement qui ne sera pas utilisable par d'autres clients souhaitant se connecter, si la limite de `max_connection` est atteinte. Pour éviter cela, il est possible d'ajuster la variable `wait_timeout`. Sa valeur par défaut est de 8 heures, ce qui est énorme. En réduisant cette valeur vous indiquerez au serveur MySQL une nouvelle limite au-delà de laquelle il coupera cette connexion si celle-ci est inactive.

Concernant les statuts des requêtes issues d'un `SHOW PROCESSLIST` (*Sending data, Writing to net, Copying to tmp table, etc.*), si certains sont évocateurs, d'autres le sont beaucoup moins. La documentation de MySQL en dresse une liste exhaustive :

▶ <http://dev.mysql.com/doc/refman/5.1/en/general-thread-states.html>

Cependant nous vous conseillons de consulter également la liste, peut-être plus lisible, de Domas Mituzas sur son blog :

▶ <http://mituzas.lt/2009/09/27/mysql-processlist-phrase-book/>

Les statuts y sont expliqués de façon beaucoup plus claire.

Si MySQL vous refuse la connexion en indiquant explicitement *Too many connections*, c'est tout simplement que le nombre défini par `max_connections` a été atteint. Il reste une chance cependant : MySQL autorise bien `max_connections` seulement à se connecter simultanément, mais il existe une connexion supplémentaire, notamment en cas d'urgence.

Cette connexion n'est disponible que pour des utilisateurs disposant du droit `SUPER` ou l'utilisateur `root`. Cet utilisateur avisé pourra alors se connecter à la base même si celle-ci a atteint son nombre maximal théorique de connexions. Rappelons quand même qu'il est déconseillé de laisser une application se connecter à la base sur le compte `root`.

ASTUCE Défilement page par page pour SHOW FULL PROCESSLIST

L'utilisation de la commande `SHOW FULL PROCESSLIST` sur un serveur chargé provoque un défilement illisible, car trop rapide, des nombreuses requêtes exécutées à cet instant sur le serveur MySQL. Il est possible de simuler le comportement du défilement page par page (le `| more` sous Linux, par exemple) grâce à la commande suivante :

```
mysql> pager more;
```

Ainsi, c'est vous qui ferez défiler à votre guise les requêtes affichées par le `SHOW FULL PROCESSLIST`. Pour retrouver le comportement par défaut du client MySQL, saisissez :

```
mysql> nopager;
```

En cas d'échec de ce dernier recours, il ne vous reste plus qu'à redémarrer le serveur MySQL pour récupérer des connexions de libres afin de vous connecter. Si le flux de connexions à la base est tel que vous craignez que la situation ne se renouvelle aussitôt une fois le serveur redémarré, sachez qu'il est possible de restreindre les connexions entrantes à celles qui sont émises localement. L'ajout de la commande `bind-address=127.0.0.1` dans le fichier de configuration `my.cnf`, permet ce type de limitation. `skip-networking` désactive l'écoute des connexions TCP/IP et n'autorise que les connexions via socket sous Unix.

REMARQUE **Se référer aux chapitres concernés**

Nous détaillons le matériel au chapitre 2 et les logs MySQL au chapitre 5 ; c'est pourquoi nous n'aborderons pas à nouveau ici le fonctionnement du RAID ou du log d'erreurs, par exemple. Il en va de même pour d'autres thèmes évoqués dans ce chapitre tels que la réplication ou la mise en place et l'exploitation d'outils de surveillance pour son système.

Supprimer les requêtes les plus lourdes

Que vous soyez l'administrateur des bases de données ou non, difficile de connaître par cœur l'intégralité des utilisateurs MySQL en activité sur les différentes bases.

La dénomination des utilisateurs MySQL rend parfois difficile leur rôle au sein du SI. De plus, certains comptes utilisateurs sont susceptibles d'être utilisés par plusieurs projets qui n'ont pas forcément de liens. Il est du coup difficile de mesurer l'importance d'une requête en la rapportant seulement au nom de l'utilisateur MySQL qui l'exécute.

En cas d'extrême urgence, il est du devoir du DBA de sacrifier certaines requêtes hautement pénalisantes pour l'ensemble du serveur MySQL afin que le site Internet reste disponible. Les internautes sont préservés au détriment de scripts internes qui devront alors être rejoués.

DANS LA VRAIE VIE

Dans un monde idéal, ces deux types d'activités ne devraient pas s'exécuter en même temps ou en tout cas pas sur le même serveur, mais la réalité économique l'emporte parfois sur les règles de sécurité.

Aussi, posséder une liste des utilisateurs dont l'activité est vitale pour l'entreprise et à ne couper sous aucun prétexte (facturation, livraison...) vous permettra de trier le moment venu. Fort de cette liste d'intouchables, les requêtes susceptibles d'être supprimées pour soulager le serveur seront plus facilement identifiables.

PRATIQUE Supprimer des requêtes rapidement

En cas d'urgence, exploiter l'affichage de la commande `SHOW FULL PROCESSLIST` en ligne de commandes via un terminal de type `putty` peut s'avérer délicat. Si le serveur est chargé votre terminal sera submergé de requêtes pour la plupart peu lisibles. Un outil tel que `mytop` permet notamment d'afficher les requêtes par couleur afin de différencier leur temps d'exécution. Il est également possible, à partir de l'interface de cet outil, de supprimer les requêtes de votre choix en indiquant leur identifiant.

MySQL Administrator est une alternative à ce type de script. Son interface graphique permet de naviguer facilement sur l'ensemble des requêtes exécutées sur le serveur. Un tri par utilisateurs MySQL est possible, le nombre d'occurrences du même utilisateur est également indiqué et un bouton permet de supprimer simplement la requête de votre choix ou toutes celles appartenant à un utilisateur.

Si vous n'utilisez pas MySQL Administrator, une alternative est de générer vous-même le fichier des identifiants à supprimer, en utilisant la base `information_schema` :

```
mysql> SELECT concat('KILL ', id, ';') FROM information_schema.processlist
      WHERE user='user_cible' INTO OUTFILE '/tmp/req2del.txt';
```

MySQL génère ainsi un fichier comportant une ligne par requête appartenant à l'utilisateur visé, comme :

```
KILL 272;
KILL 284;
[...]
```

Il ne reste plus qu'à l'exécuter :

```
mysql> source /tmp/req2del.txt;
```

ou

```
shell> mysql -uuser -p < /tmp/req2del.txt
```

À noter : même détruite, une requête peut prendre un long moment avant de rendre la main. En effet, un `ROLLBACK` d'une grosse opération sous InnoDB peut être très long, il n'y a pas de solution idéale dans ce cas là... Si vous redémarrez le serveur vous vous exposez à une récupération de crash d'InnoDB qui peut également prendre un certain temps.

Autre solution extrême à réserver à de gros `SELECT` bloquants sous MyISAM, par exemple : récupérer l'identifiant du thread correspondant à la connexion qui exécute la requête et effectuer un `kill -9`. À vous de mesurer les conséquences d'une telle coupure ; avec MyISAM, il n'y a pas de notions de transactions et la cohérence de la base est donc peut-être compromise à ce niveau.

Encore une fois, testez ces outils par vous-même avant qu'une situation d'urgence ne se déclare. Nos conseils sur la surveillance de vos serveurs occupent le chapitre 4.

Il suffit parfois de supprimer une requête pour débloquer un empilement de connexions préjudiciable pour le serveur MySQL. C'est le cas si cette requête n'est pas optimisée et génère un trop grand nombre de verrous, créant ainsi une file d'attente importante. MyISAM et son verrouillage par table est plus impacté qu'InnoDB mais aucun de ces moteurs de stockage n'est à l'abri de tels phénomènes.

Éviter que l'authentification des utilisateurs repose sur un DNS : l'erreur `unauthenticated user`

Difficile de supprimer toutes les requêtes d'un utilisateur particulier si même MySQL ne sait pas l'identifier... Ce type de message « `unauthenticated user` » est visible en cas de serveur MySQL très chargé ou en proie à un problème de DNS. Nous vous recommandons de ne pas les utiliser lorsque vous définissez un utilisateur MySQL.

Définition d'un utilisateur MySQL basé sur un DNS

```
mysql> GRANT SELECT on *.* TO 'me_dns'@'dev.example.org' IDENTIFIED BY 'password';
```

Le même exemple sans l'utilisation de DNS

```
mysql> GRANT SELECT on *.* TO 'me'@'192.168.10.11' IDENTIFIED BY 'password';
```

En cas de problème du serveur DNS, MySQL ne pourrait identifier le premier utilisateur `'me_dns'`, le qualifiant de `unauthenticated user`. Si vous rencontrez ce problème, il est probable que vous en venez à bout en effectuant ces deux actions :

- bannir les DNS de la définition de vos utilisateurs ;
- démarrer le serveur avec l'option `--skip-name-resolve` (nous vous conseillons d'activer cette option).

Consulter son système de surveillance

Une situation d'urgence ? Voici le bon moment pour capitaliser sur vos efforts précédents et consulter vos logs ou graphiques d'activité liés à la base.

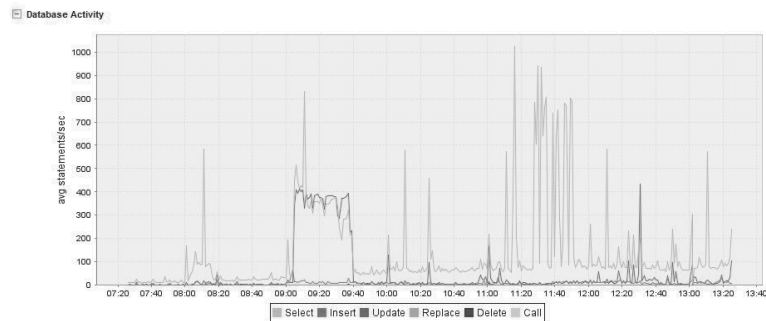
Un graphique est souvent bien plus explicite que différentes métriques à analyser. De plus, selon le degré d'historisation que vous utilisez, il est sûrement possible de remonter (pourquoi pas) à l'heure d'origine du problème. Cette dernière information est cruciale pour incriminer un script programmé dont les impacts ne sont pas forcément simultanés avec son heure de lancement. Pensez par exemple au résultat d'une jointure complexe qui serait stocké par le langage de script afin de générer une boucle de mises à jour mal optimisée, le tout sur une table très volumineuse... L'effet différé est garanti et potentiellement dramatique.

Afin de déterminer si le problème rencontré sur le serveur provient d'une requête ponctuelle passée directement via un client MySQL en production (gestion des droits à revoir !) ou est issu d'un script programmé, un ordonnanceur permet d'identifier rapidement quel script tourne à quelle heure. Son utilisation est moins fastidieuse que le parcours des crontab mais encore faut-il en posséder un. Une fois déter-

minée l'heure probable de l'origine du problème, comparez celle-ci avec celles présentées par l'ordonnanceur.

Figure 1-1

Posséder un historique sur l'activité des bases est d'une grande utilité, exemple avec MySQL Enterprise Manager.



Tranche de vie d'une campagne marketing improvisée

Votre système de surveillance est en alerte, les diodes rouges remplacent frénétiquement leurs paisibles ancêtres de couleur verte bien plus rassurant. Pourquoi ? Les tâches programmées ne sont pas incriminées : les crontab ont été passées au peigne fin. Les développeurs vous jurent qu'ils n'y sont pour rien. Finalement, une fois la tempête passée, quelqu'un a déclenché le lancement d'une campagne marketing reposant sur des pop-ups dynamiques aux requêtes non optimisées. Outre le problème de communication entre les départements technique et marketing, l'incident révèle qu'un seul et même utilisateur peut à lui seul occuper 100 % des connexions disponibles sur votre serveur MySQL.

Afin d'empêcher qu'un seul utilisateur consomme toutes les ressources, il est possible d'ajouter certaines restrictions par heure à cet utilisateur. Ainsi, le nombre de requêtes, de mises à jour ou de connexions, sont des limites qu'on peut fixer. De plus, le nombre maximal de connexions simultanées pour un utilisateur est lui aussi configurable.

Restriction à 200 connexions simultanées maximum pour un utilisateur existant :

```
mysql> GRANT USAGE ON *.* TO 'user_pub'@'192.168.200.20'  
      with MAX_USER_CONNECTIONS 200;  
mysql> SHOW GRANTS FOR 'user_pub'@'192.168.200.20';  
Grants for user_pub@192.168.200.20: GRANT SELECT ON *.* TO  
'user_pub'@'192.168.200.20' IDENTIFIED BY PASSWORD  
'*1270C0C06DEE42FD1618BB99005ADCA2EC9D1E19'  
      WITH MAX_USER_CONNECTIONS 200
```

Une augmentation brutale du nombre de requêtes peut provenir d'un problème du cache : soit celui-ci a été brutalement désactivé ou invalidé (query cache) soit votre système basé sur memcached est défectueux.

Temps de résolution : une heure

Estimer que l'on a au moins une heure devant soi signifie probablement que le problème ou ses conséquences est en partie identifié. C'est souvent le cas d'une surcharge soudaine sur le serveur MySQL. Attribuée à un script programmé, il est possible de retrouver une situation normale en coupant le programme concerné. Le but est ici de ne pas couper brutalement le responsable mais plutôt de l'optimiser.

Si nous disposons d'une heure pour résoudre le problème, cela laisse le temps d'optimiser une requête complexe et coûteuse, de diffuser le correctif aux équipes concernées et d'assister peut-être à une mise en production dans l'heure si le problème survient en heures ouvrables.

La chasse aux requêtes lentes

Dans l'hypothèse où notre problème soit issu d'une ou plusieurs requêtes trop coûteuses, le but est tout d'abord de les identifier rapidement puis de les analyser afin de les optimiser. Si votre outil de surveillance préféré (MySQL Administrator, mytop, etc.) est suffisant pour repérer les requêtes les plus lourdes s'exécutant en temps réel sur le serveur, ils ne sont d'aucune utilité pour analyser le passé.

En effet, afin de retrouver les requêtes qui ont pu impacter votre serveur et qui sont désormais peut-être terminées (mais appelées à s'exécuter de nouveau), des logs sont nécessaires. MySQL possède un fichier de logs dédiés aux requêtes lentes : il s'agit du `slow query log`. Celui-ci peut par exemple être exploité par `mysqldumpslow` ou `mysqlsa`.

Ne recherchez pas uniquement les requêtes les plus lentes mais également les plus fréquentes. En effet, votre charge MySQL est susceptible de provenir d'une grande quantité de petites requêtes n'apparaissant pas dans les logs de requêtes lentes selon le critère utilisé le plus souvent : le temps d'exécution. 100 000 requêtes de 0,9 seconde chacune n'apparaissent pas si le `long_query_time` est défini à une seconde ; pourtant, celles-ci sont peut-être responsables d'une surcharge serveur.

À LIRE **Un chapitre consacré à l'étude des journaux**

Au sujet des logs, lire ou relire notre chapitre 5 consacré aux logs MySQL. Il est possible d'agir sur la façon dont le serveur MySQL génère les journaux des erreurs, de requêtes lentes ou généraux. Il est précieux, tout particulièrement en situation d'urgence, de connaître leur fonctionnement.

Réécrire les requêtes trop coûteuses

Une heure, cela laisse le temps d'agir sur un nombre restreint de requêtes un peu trop gourmandes. Cependant le temps passe vite : certaines requêtes nécessitent à elles

seules ce laps de temps pour améliorer les choses. Il est parfois difficile d'analyser le bien-fondé fonctionnel d'une requête complexe multipliant les jointures et les sous-requêtes. Aussi, si cela est possible, proposez au développeur de se joindre à vous pour une première analyse de celles-ci. Les critères de jointures sont-ils toujours pertinents ?

Certains développeurs sont capables de créer du code SQL complexe et valide mais c'est également à l'administrateur de bases de données de leur signaler le fonctionnement interne de MySQL et d'évoquer certaines faiblesses propres à la branche 5.x : sous-requêtes imbriquées et/ou corrélées aux requêtes externes. Désormais au courant de nouvelles bonnes pratiques, le développeur est susceptible de corriger lui-même certaines des requêtes remontées lors de votre recherche de requêtes lentes, vous laissant ainsi du temps disponible pour en corriger d'autres. Une fois la correction du développeur effectuée, comparez les `EXPLAIN` avant et après et classez l'affaire si tout va bien.

Nous attirons votre attention sur le fait que mettre en production une requête sans passer par une phase de test conforme à celle que vous effectuez d'habitude est une source de régression potentielle. Une situation urgente autorise certains raccourcis susceptibles de vous sortir d'affaire, mais il n'est pas pour autant conseillé d'oublier toute prudence ; ayez au moins à l'esprit le risque lié à cette pratique.

Nos conseils pour optimiser vos requêtes se situent dans le chapitre 6, consacré à l'optimisation de la base de données.

ASTUCE Les tables statiques à la rescousse

Si vous estimez qu'il est impossible en une heure de venir à bout de toutes les requêtes à corriger, pensez aux tables statiques ou *summarized tables*. Cela ne fonctionne pas dans tous les cas mais cette technique est très efficace si elle convient. L'idée est de générer une fois pour toutes les résultats d'une requête complète et de les insérer dans une autre table. Plutôt que d'exécuter n fois une requête complexe et lente, elle est jouée une seule fois et ses résultats archivés. L'inconvénient, bien sûr, est que le contenu de la table statique ne sera pas mis à jour. À vous donc de définir éventuellement la périodicité de cette opération (vidage de la table et nouvelles insertions à partir de la requête complexe), un script programmé (`cron`) peut s'en charger. Il est également possible d'utiliser les `event` au sein même de MySQL, vous en trouverez une description sur l'un de nos blogs :

▶ <http://dasini.net/blog/2009/06/16/>

Cette technique a un sens lorsqu'une requête ne peut être mise en cache très longtemps (résultats invalidés par MySQL si par exemple l'une des tables impliquées par la requête est mise à jour). Il est également nécessaire que cette fameuse requête complexe ne soit pas contextuelle à la session d'un internaute. Dans ce dernier cas, les tables temporaires peuvent constituer une bonne solution à condition que leur taille maximale ne soit pas un problème (il s'agit de la plus petite valeur entre ces deux variables système : `max_heap_table_size` et `tmp_table_size`). Dans ce cas la mise en cache dans une table statique n'aurait pas de sens, le contenu doit servir au plus grand nombre. Il peut donc s'agir d'un contenu dynamique dont on peut se permettre une fréquence de rafraîchissement inférieure au temps réel. Un exemple : statistiques et calculs autour des points d'un championnat quelconque. Les résultats ne seront pas obsolètes avant la prochaine épreuve.

Les problèmes de réplication

Le chapitre 8, consacré à la réplication, détaille les mécanismes, la configuration et les commandes nécessaires pour gérer la réplication MySQL. Il est important que son fonctionnement soit bien compris.

En conséquence, nous n'expliquerons pas ici à nouveau les arcanes de la réplication. En revanche, le conseil à retenir dans le cadre d'une situation d'urgence concernant une réplication est de ne pas agir à la légère. En effet, en cas d'erreur de votre part, par exemple lors du redémarrage d'une réplication arrêtée, vous pouvez rendre l'esclave inutilisable avec pour seule issue de recharger complètement les données qu'il contient. L'importance d'un tel incident (esclave hors service) est inversement proportionnelle au nombre d'esclaves possédés. C'est une situation dramatique si vous n'en possédez qu'un seul (faire pointer les applications vers le maître si celui-ci tient la charge...) mais moins grave si vous en possédez un grand nombre. L'incident pourrait passer inaperçu si votre système est correctement dimensionné avec une bonne répartition de charge.

Les messages d'erreurs liés à la réplication sont assez variés. Ne pas comprendre l'origine du problème vous expose à de graves difficultés. Il serait inconsidéré d'appliquer à la légère le fameux `SET GLOBAL sql_slave_skip_counter=1` indiquant à l'esclave de passer outre le problème qui a généré l'erreur et de sauter à l'instruction suivante du `relay-log`. Le risque de désynchronisation entre l'esclave et le maître serait alors important et d'autres erreurs analogues ne manqueraient pas de se produire à plus ou moins courte échéance.

Pire encore, ignorer systématiquement un certain type d'erreurs (`slave-skip-errors`) sans maîtriser réellement les conséquences de ce réglage, est une grosse erreur.

Selon le type de moteur de stockage que vous utilisez (MyISAM ou InnoDB), et le type de réplication (`row based` ou `statement based`) une requête stoppée manuellement sur le maître peut avoir des conséquences différentes sur l'esclave. Avec MyISAM et en mode `statement based`, par exemple, une requête interrompue sur le maître provoque :

```
Query partially completed on the master (error on master: 1053) and was
aborted. There is a chance that your master is inconsistent at this
point. If you are sure that your master is ok, run this query manually
on the slave and then restart the slave with SET GLOBAL
SQL_SLAVE_SKIP_COUNTER=1; START SLAVE;
```

L'esclave détecte que la requête a été interrompue sur le maître (cette information circule dans le log binaire) et décide de stopper lui aussi la réplication. À vous de vérifier que la ou les tables concernées sur le maître et l'esclave sont bien symétriques : il n'y a

pas de `ROLLBACK` sur `MyISAM`. Si c'est le cas, vous pouvez alors utiliser le conseil prodigué par le message d'erreur `SET GLOBAL...` et redémarrer la réplication.

La réplication `MySQL` est asynchrone : une information enregistrée sur le maître n'est pas immédiatement disponible pour l'instruction suivante sur tous les esclaves. En conséquence, certaines applications ou scripts rencontrent des problèmes. C'est le cas typique cité précédemment : une insertion est jouée sur le maître et le script recherche immédiatement son identifiant correspondant sur l'esclave ; s'il ne le trouve pas, son comportement peut s'en trouver perturbé.

Une façon simple de gérer ce souci est d'effectuer les requêtes de ce type (récupération d'une clé liée à l'insertion d'un enregistrement) sur le maître.

Temps de résolution : une journée

À l'échelle de la journée nous sommes pour ainsi dire sortis de l'urgence. La survie de la base n'est plus une question de minutes mais plutôt de jours. Cependant, le scénario que nous suivons (rappelez-vous du découpage arbitraire en trois parties) implique que la situation est préoccupante et demande des restructurations parfois importantes.

Une journée ne représente que quelques heures durant lesquelles il va falloir trouver des solutions. À ce stade le problème est en effet détecté et il s'agit probablement des prémices d'une montée en charge difficile. Peut-être rencontrez-vous de façon récurrente des ralentissements à horaires réguliers, des requêtes de plus en plus lentes, des tables sur lesquelles il est difficile d'effectuer des opérations de maintenance sans bloquer la production.

LIRE Chapitre 8 consacré à la réplication

Les indisponibilités de la base liées aux changements de structures `ALTER`, `DROP`, ou de versions peuvent être résolus en partie grâce à la réplication.

Bref, il est temps de prendre le problème à bras-le-corps et d'y consacrer du temps à l'avenir puisque c'est finalement cela dont il s'agit : anticiper les problèmes de demain.

Face à des problématiques d'une telle envergure, les réponses techniques ne sont pas forcément immédiates, tout du moins pas applicables en pleine journée. Certaines phases sont en effet bloquantes : modification de tables (`ALTER TABLE`), refonte d'une partie de l'application (phase de mise en production), remaniement des crons... tout cela mérite quelques tests. Cependant si une opération coup de poing est décrétée sur un thème précis, comme résoudre les problèmes récurrents de performance du serveur xxx, il s'agit d'agir vite et bien. Toutes les difficultés ne seront probablement pas

réglées le soir même mais les solutions à appliquer et les premiers correctifs sont applicables pour le lendemain (les premiers scripts correctifs sont susceptibles de s'exécuter pendant la nuit si votre activité s'y prête).

PRATIQUE **Un problème peut en cacher un autre**

Sur une telle opération, il est souhaitable d'impliquer les utilisateurs concernés par les modifications de base. Il est fréquent qu'en discutant d'un problème de performance les utilisateurs fassent part au DBA qu'ils ont par ailleurs des retours de blocage intermittents à différents horaires, des pertes de connexion (*MySQL Server has gone away*), etc. Plutôt que d'effectuer plusieurs réunions pour différents problèmes sur un même serveur, autant tous les évoquer au moins une fois et tenter de les résoudre en même temps, si possible. Un passage en InnoDB peut résoudre les problèmes de verrous évoqués par les utilisateurs ; un index contribuera à accélérer certaines requêtes ainsi que le partitionnement d'une table volumineuse. De plus, si la réunion n'est pas purement orientée bases de données, un administrateur système ou réseau sera le bienvenu pour régler certains problèmes de connexion ou de perte de paquets, constatés éventuellement sur certaines machines.

Afin de dessiner une solution face à une montée en charge avérée ou à venir, un passage par le chapitre 4, consacré à la surveillance de votre serveur, est tout indiqué. En fonction des conclusions de cet audit, il sera possible d'émettre un planning et de classer les actions à effectuer par priorité ou faisabilité.

Voici un échéancier tel que l'administrateur de bases de données pourrait l'écrire à l'issue d'une telle réunion :

- lundi :
 - mise à jour matérielle nécessaire (+8 Go RAM et 2 quad-cœurs, attente des composants) ;
 - optimisation des requêtes facturation ;
 - partitionnement de la table « factures » sur l'environnement de développement ;
 - benchmarks de la version partitionnée (mysqslap) ;
- nuit lundi/mardi :
 - ALTER TABLE factures (partitions) si tests concluants ;
 - passer les tables factures_client_xx en InnoDB ;
- mardi :
 - consulter logs des actions ALTER partitions/passage en InnoDB ;
 - surveiller cron 10h45 et comparer situation de la veille.

Conseils généraux face à l'urgence

Tirer profit du passé

Documenter les commandes ou procédures clés à appliquer en cas de problème au travers d'un wiki, d'un fichier texte, etc., permet le jour *J*, quelle que soit l'heure, d'être opérationnel. Pensez à votre état de fraîcheur si l'on vous réveille lors d'une astreinte à 3h du matin... Ne pas perdre de temps à se rappeler comment éteindre ou rallumer certains services importants en plus des procédures liées à l'activité MySQL n'est pas négligeable dans pareilles situations.

Si ce n'est pour vous, faites-le pour les personnes susceptibles de prendre le relais en votre absence. Ce document doit être enrichi régulièrement, au fil des problèmes rencontrés, des changements de serveurs, de DNS, etc.

Anticiper les problèmes

C'est évident mais tellement vrai : mieux vaut prévenir que guérir. Lorsque les problèmes surviennent il est parfois déjà trop tard. Outre les urgences quotidiennes, l'administrateur de bases de données doit veiller à se ménager du temps, chaque jour dans l'idéal, pour améliorer l'existant, effectuer quelques opérations de maintenance (`OPTIMIZE TABLE`, `ANALYZE TABLE`, `CHECK TABLE` ou sa version client bien pratique `mysqlcheck`), repérer les tables volumineuses susceptibles de poser problème (partitionnement nécessaire ?), suivre de près les serveurs du moment, ceux dont l'activité est particulièrement forte ou importante à un instant précis.

ATTENTION Modifications à chaud, en production

Soyez extrêmement vigilant à l'égard des modifications qui sont susceptibles d'être exécutées à chaud sur les machines de production. Certaines variables système sont en effet modifiables sans redémarrage du serveur MySQL. Il est important d'en mesurer tous les impacts possibles en termes de performance et de stabilité du système. Il est possible de mettre un serveur à plat en modifiant une seule ligne dans un fichier de configuration `my.cnf` ; par exemple, passer `sync_binlog` de 0 à 1 peut entraîner une chute dramatique des performances de tout le système disque (au sujet de cette variable, lire également le chapitre 8, consacré à la réplication, et le 2, dédié au matériel). De manière générale, il faut reporter dans le fichier de configuration de MySQL toute modification effectuée (et validée) à chaud. Sans cela, cette nouvelle valeur serait perdue au prochain redémarrage du serveur.

Quelques bonnes pratiques pour terminer :

1. archivez dans un document toute modification effectuée ainsi, afin qu'en cas de problème vous puissiez remonter le fil de vos actions. Notez les horaires ;
2. reportez ces modifications dans le `my.cnf` ;
3. envoyez par e-mail à votre équipe les modifications effectuées. L'équipe d'astreinte ou les administrateurs système pourront ainsi éventuellement mettre en relation un problème et votre modification, accélérant la résolution du cas.

L'entraînement à l'urgence

Il ne suffit pas d'installer un maître et deux esclaves pour se croire à l'abri de tout problème... Que se passe-t-il en cas de perte du master ? Dans un autre registre, les sauvegardes sont-elles fiables ? On lit souvent qu'il faut tester ses sauvegardes, c'est-à-dire les recharger sur un serveur et s'assurer que les données soient bien présentes et correctes ; mais combien d'entre nous l'effectuons réellement ? Si cela devait se produire en production, en combien de temps seriez-vous capable de tout remonter ? Ces types de scénarios catastrophes se produiront un jour ou l'autre et il faut y être préparé. Si les machines vous manquent pour tester tout cela, il est possible d'utiliser des solutions telles que MySQL Sandbox pour se construire une configuration analogue à celle de la production à moindre coût. Autre test intéressant à effectuer sur des serveurs de test : débrancher physiquement le courant et rallumer le tout. Où en est la réplication ? Des tables sont-elles corrompues ? En combien de temps le serveur MySQL redémarre-t-il ? Difficile de le deviner : il faut tester. L'utilisation de machines virtuelles permet également de se constituer un parc de machines avec lesquelles s'entraîner.

Enregistrer les données de l'incident

Quel que soit le degré d'urgence dans lequel vous vous trouvez, à des fins immédiates ou pour plus tard, tentez de sauvegarder des informations concernant l'état du serveur au moment de l'incident. Le chapitre 4, consacré aux solutions de surveillance, est un bon point d'entrée.

En quelques mots, archiver les résultats de commandes systèmes telles que :

```
iostat -dx 5
vmstat 1 10
top -b -n 1
cat /proc/meminfo
```

et du côté de MySQL (si cela est possible) :

```
SHOW FULL PROCESSLIST;
SHOW INNODB STATUS\G
SHOW GLOBAL STATUS;
```

L'état d'esprit à adopter dans l'urgence

Terminons ce chapitre sur quelques règles élémentaires en cas de coup dur.

- Ne paniquez pas, restez concentré.

- Tentez de mesurer rapidement l'impact du problème sur le reste du système/des applications.
- Attention aux conclusions hâtives.
- Soyez méthodique.
- Connaître le fonctionnement des applications est un plus. Sans autres ressources humaines que vous (en pleine nuit par exemple), vous pourrez agir sur certains scripts ou identifier plus rapidement la source du problème.
- Dans l'urgence, allez au plus simple.
- Si possible, effectuez une sauvegarde de ce que vous manipulez, un service dégradé vaut mieux que pas de service du tout.

Trouver de l'aide

Vous êtes dépassé par le problème ? Cela arrive. Dans l'intérêt de tous, ne restez pas seul à tenter de résoudre l'affaire si vous n'y arrivez pas et n'avez aucune piste. Si aucune aide n'est disponible en interne et que vous tournez en rond, il est temps de solliciter une aide extérieure. Lisez le chapitre 8 consacré à ce sujet. Il existe des services de support joignables par Internet ou par téléphone.