



Guide pratique de la migration vers MySQL 5.0



Livre blanc MySQL®
Juillet 2006

Sommaire

Introduction	3
Questions économiques de la migration vers MySQL.....	3
Évaluer le coût d'utilisation de MySQL.....	3
Un rapide coup d'œil à MySQL AB.....	4
Valider les capacités techniques de MySQL	5
Examen des performances de MySQL.....	6
MySQL, la haute fiabilité	7
Un logiciel simple mais riche en fonctions	8
Suggestions pratiques pour une migration facile vers MySQL	9
Étape 1 – Documenter les sources	10
Étape 2 – Concevoir les cibles	10
Étape 3 – Concevoir les flux ETL	11
Étape 4 – Construire/effectuer des migrations	12
Étape 5 –Tests de performance	12
Conclusion.....	14
Ressources.....	14

Introduction

Depuis la sortie récente de MySQL 5.0 et sa rapide adoption par le marché, de nombreuses entreprises, administrations, établissements d'enseignement et autres entités songent à quitter leurs coûteux fournisseurs de bases de données propriétaires pour migrer vers une plate-forme open source. Naturellement, migrer une base de données, quelle qu'elle soit, n'est pas une mince affaire et beaucoup s'interrogent sur les éléments à prendre en considération pour migrer vers MySQL. Des questions comme « le modèle open source est-il réellement moins coûteux à long terme ? » ou « Est-ce que MySQL est adapté aux besoins de mes applications stratégiques ? », entre autres, se posent tout naturellement.

Ce document tente d'apporter une réponse à ces interrogations ainsi qu'aux autres questions clés qui se posent lors de la migration d'une application stratégique vers MySQL. La migration vers MySQL sera abordée sous ses aspects à la fois économiques et techniques. Que vous soyez un manager ou un DBA chevronné, vous y trouverez donc les réponses aux questions que vous vous posez concernant la migration vers la base de données open source la plus populaire au monde : MySQL.

Questions économiques de la migration vers MySQL

Avant de choisir une technologie, les entreprises modernes doivent commencer par évaluer les raisons économiques qui justifient son acquisition, au-delà des promesses des vendeurs de logiciels. Bien que chaque entreprise suive sa propre méthodologie d'évaluation, les facteurs clés sont en général les suivants :

1. Le calcul du coût total de possession – Le solde financier de l'opération est ce qui retient le plus souvent l'attention. Pourtant, il est parfois difficile de vérifier le réel coût de l'investissement dans une plate-forme. Les principaux postes qui entrent dans ce calcul sont généralement les coûts de licence, de maintenance et de support, mais d'autres aspects devraient également y figurer, comme la formation du personnel et les coûts en matériel.

2. Valider la viabilité de l'éditeur du logiciel – Les entreprises souhaitent s'assurer de la viabilité de leur fournisseur. Il est donc important d'analyser sa situation économique, par exemple en vérifiant sa santé financière actuelle, le taux de pénétration de son logiciel sur le marché et sa liste de clients.

Examinons plus en détail les deux points que nous venons d'évoquer : le coût total de possession de MySQL, et la santé de MySQL AB.

Évaluer le coût d'utilisation de MySQL

L'une des principales raisons de l'énorme popularité des logiciels open source sur le marché professionnel actuel est l'économie considérable qu'ils permettent de réaliser. On sait qu'un grand nombre des entreprises du classement Global 2000 dépensent de 500,000 à 10 millions de dollars par an en achat de nouvelles licences et en coûts de maintenance pour leurs logiciels de base de données propriétaires existants. Pour la plupart d'entre elles, MySQL constitue donc une alternative incroyablement attractive lorsqu'elles constatent qu'elles peuvent diminuer leurs coûts de 80% à 90 % en utilisant MySQL à la fois sur de nouveaux projets d'application, et en faisant évoluer leurs systèmes existants pour utiliser MySQL au lieu de solutions plus coûteuses.

De telles économies de licence et de maintenance sont encore amplifiées lorsqu'il s'agit d'entreprises contraintes de faire évoluer leurs applications pour répondre à une plus forte demande de leurs utilisateurs. Par exemple, une récente étude publiée par Charlie Garry, une autorité en matière de bases de données, compare les coûts de déploiement d'Oracle RAC et de MySQL et il exprime l'opinion suivante :

« Oracle a encouragé ses clients à évoluer avec son approche RAC. Bien qu'il y ait techniquement du pour et du contre à utiliser RAC, il importe de prendre en compte le coût supplémentaire. Puisque RAC fait l'objet d'un surcoût et que la majorité des utilisateurs de RAC acquièrent aussi la fonction de partitionnement de données pour faciliter l'évolution, le coût par processeur additionnel est de 85 400 \$, frais de licences, support et mises à jour obligatoires compris. Si nous mettons cela en perspective, pour le prix de chaque processeur ajouté, vous pourriez acheter environ 11 serveurs d'entrée de gamme, installer une base de données open source telle que MySQL et vous offrir un support technique Gold pour la base de données pour chacun des serveurs ¹ ».

Alors que les coûts de licence et de maintenance/support représentent l'essentiel du coût total de possession, les dirigeants doivent également prendre en compte les frais annexes de matériel et de formation du personnel souvent occasionnés par une migration de logiciel.

Grâce à l'architecture de MySQL, la nécessité de faire évoluer le matériel ne se justifie plus du tout. C'est même le contraire en général car MySQL fonctionne plus efficacement sur du matériel Intel/Opteron d'entrée de gamme que les bases de données propriétaires. De nombreux utilisateurs de MySQL ont donc la possibilité de réduire leurs dépenses en matériel par la même occasion. De même, nous constatons une adoption significative de MySQL dans le secteur de l'enseignement, où l'on doit souvent se contenter de machines plus anciennes.

Les coûts de formation du personnel sont facilement maîtrisés lors du passage à MySQL, car la base de données est simple à utiliser et à comprendre. Par exemple, les professionnels des bases de données qui viennent d'Oracle ou de Microsoft SQL Server n'auront aucun problème à apprendre MySQL, dont l'architecture et les fonctionnalités correspondent étroitement à ces bases de données. D'ailleurs, certains les abandonnent au plus vite, comme l'a révélé un récent article de silicon.com :

« Paul Broome, Directeur informatique de 192.com, annonce son intention de migrer le plus tôt possible de Windows server et de SQL server. Ainsi qu'il l'explique : « C'est un véritable boulet pour une PME. Pour nous permettre de croître, tous nos nouveaux serveurs de bases de données vont utiliser Linux 64 bits et MySQL. Nous en avons assez. Nous ne pouvons plus évoluer avec Microsoft, c'est trop cher ».²

Un rapide coup d'œil à MySQL AB

Migrer vers une nouvelle plate-forme de base de données signifie en général s'engager sur le long terme. Il est donc naturel que les entreprises modernes qui envisagent d'évoluer vers MySQL veuillent connaître l'historique de la base de données, qui l'utilise, et comment.

Selon les études réalisées dans l'industrie informatique, la première inquiétude des entreprises qui envisagent d'adopter un logiciel open source est le manque d'assistance.³ Une grande société ne peut évidemment mettre en place pour ses applications névralgiques qu'un logiciel bénéficiant du soutien d'un éditeur reconnu, d'une excellente réputation et offrant un support de première classe.

¹ Database Scale-Out Using Open Source Software: Achieving Infrastructure Agility, septembre 2005.

² <http://www.silicon.com/ciojury/0,3800003161,39152500,00.htm>, septembre 2005.

³ Forrester Research, février 2004

De nombreuses entreprises adoptant MySQL pour la première fois sont étonnées d'apprendre que MySQL AB existe depuis plus de dix ans. Travaillant dans plus de 25 pays et comptant plus de dix millions d'installations, MySQL AB poursuit son expansion et est de loin la base de données open source la plus populaire au monde. Chaque jour, plus de 50 000 personnes sur la planète téléchargent le logiciel MySQL pour faire fonctionner leurs applications professionnelles, ce qui explique que MySQL ait doublé son chiffre d'affaires chaque année depuis trois ans.

En 2005, la popularité de MySQL a continué de progresser, comme l'a montré Evans Data en signalant la multiplication des utilisations de MySQL :

« Les déploiements de base de données open source ont augmenté de plus de 20 % ces six derniers mois, selon une étude d'Evans Data effectuée à l'automne 2005 sur le développement des bases de données (Database Development Survey). L'utilisation de MySQL, par exemple, a augmenté de plus de 25 % en six mois et devient presque majoritaire dans le monde des bases de données, 44 % des développeurs utilisant cette base de données open source ⁴ ».

Devant l'accroissement de sa clientèle d'entreprises, MySQL offre désormais MySQL Network aux clients qui souhaitent un support et des services professionnels de premier ordre. À ces entreprises, MySQL Network propose un logiciel professionnel certifié, une assistance 24 h/24, des conseillers en logiciel, une politique d'indemnisation et bien plus encore. De grands fournisseurs internationaux de solutions et de systèmes informatiques ont reconnu la popularité de MySQL Network et ont passé un accord de partenariat avec MySQL pour revendre MySQL Network à leur vaste base de clients. Parmi ces revendeurs partenaires de MySQL figurent désormais HP, Novell, Dell, Unisys et d'autres vendeurs.

Pour ceux qui souhaitent obtenir des références et s'assurer de la fiabilité de MySQL vis-à-vis de leurs applications stratégiques, il suffit de regarder la liste de ses utilisateurs, qui fait véritablement figure de « Who's-Who » de la profession, et de constater la rapidité avec laquelle elle croît. Des organisations modernes telles que Yahoo, Google, Alcatel, Nokia, Franprix & Leader Price, Sabre Holdings (travelocity), Amazon, la NASA et bien d'autres, dans tous les secteurs d'activité, figurent dans le fichier clients de MySQL. Ces sociétés, comme Weather Channel par exemple, considèrent MySQL comme leur meilleur atout dans leur quête de performances élevées et d'un coût total de possession réduit pour leur base de données :

« À l'heure actuelle, Weather.com est le dixième site Web le plus important du monde et s'appuie presque à 100 % sur la base de données open-source MySQL. « En passant d'un environnement Sun Solaris à un environnement Linux sur Intel, en supprimant les frais de maintenance matérielle, en remplaçant un logiciel commercial par un logiciel open source, nous avons réduit nos coûts d'un tiers et augmenté les capacités de traitement du site Web de 30 % », indique Dan Agronow, Directeur Technique de Weather Channel Interactive ⁵ ».

Valider les capacités techniques de MySQL

MySQL ayant brillamment réussi le test décisif de l'examen économique, il lui reste à passer l'examen technique pour savoir s'il est un bon candidat à la migration. Bien que chaque entreprise possède ses propres critères techniques pour valider un logiciel, les logiciels de base de données en particulier peuvent presque toujours être évalués en termes de performances, de fiabilité et de facilité d'utilisation.

⁴ http://www.evansdata.com/n2/pr/releases/EDCDB05_02.shtml

⁵ <http://www.cio.com/archive/090105/forum.html>, septembre 2005

Examen des performances de MySQL

Les applications à haut niveau de visibilité requièrent des performances élevées. Pour certains systèmes, des performances médiocres peuvent avoir des conséquences aussi graves qu'une panne si l'on prend en considération la perte de clients ou l'interruption réelle de l'activité.

L'une des caractéristiques de MySQL est sa performance exceptionnelle dans tous les domaines d'activité des bases de données, qu'il s'agisse de traitement des transactions, de l'entreposage de données ou de sites Web à fort trafic. De nombreuses entreprises modernes obtiennent des gains de performance incroyables en utilisant une architecture évolutive (scale-out) couplée à MySQL. Dans une telle configuration, les données sont répliquées sur plusieurs serveurs MySQL (serveurs standards du marché) et la charge est répartie sur l'ensemble des serveurs. De ce fait même les requêtes le plus lourdes peuvent être exécutées avec un temps de réponse phénoménal.

MySQL dispose de fonctionnalités qui permettent de tels niveaux de performances. Par exemple, une structure originale de mise en mémoire cache offre des temps de réponse quasi instantanés pour des requêtes effectuées de façon identique. D'autres caractéristiques, telles qu'un optimiseur de base de données haut débit, de puissantes fonctions d'indexation (y compris en texte intégral), des utilitaires de chargement rapide des données, etc., aident MySQL à assurer les performances requises par des applications à haute visibilité.

Ces caractéristiques permettent à MySQL de se placer en tête des tests de performance du secteur, notamment ceux réalisés par la Standard Performance Evaluation Corporation (SPEC). Dans une récente évaluation de performances, MySQL a réalisé les meilleurs résultats à la fois pour le total des opérations effectuées et pour le coût total des opérations.⁶

Rapport prix/performances

(Les chiffres les plus bas indiquent un meilleur rapport prix/performances)

Base de données	Serveur d'appli.	Matériel de la base de données	Matériel du serveur d'appli.	\$ US/MAX.
MySQL	BEA WebLogic	Sun Fire V40z	Sun Fire V20z	82,74 \$
SQL Server	BEA WebLogic	Unisys ES7000 Aries 410	Unisys ES3020L Server	139,84 \$
Oracle	Oracle	HP ProLiant DL580 G2	HP ProLiant ML370 G3	150,68 \$
SQL Server	BEA WebLogic	Unisys ES7000 Aries 420	Unisys ES3020L Server	157,66 \$
Oracle	Oracle	HP ProLiant DL580 G2	HP ProLiant ML370 G3	168,62 \$
SQL Server	BEA WebLogic	HP ProLiant DL580 G2	HP ProLiant DL580 G3	200,34 \$
IBM DB2	IBM Websphere	IBM eServer pSeries p650	IBM eServer xSeries x335	330,07 \$

Figure 1 - Le rapport prix/performances de MySQL mesuré par la SPEC (2004).

⁶ SPECjAppServer Benchmark, SUN, BEA, MySQL, septembre 2004.

Performances des entrées avec des bases de données fonctionnant sur du matériel à 4 processeurs
(Les chiffres les plus élevés indiquent les performances les plus rapides)

Base de données	Serveur d'appli.	Matériel de la base de données	Matériel du serveur d'appli.	MAX.
MySQL	BEA WebLogic	Sun Fire V40z	Sun Fire V20z	1 363,88
Oracle	Oracle	HP ProLiant DL580 G2	HP ProLiant ML370 G3	1 165,06
SQL Server	BEA WebLogic	HP ProLiant DL580 G2	HP ProLiant DL580 G3	1 037,02
Oracle	Oracle	HP ProLiant DL580 G2	HP ProLiant ML370 G3	863,40

Figure 2 - le total des opérations par seconde surpasse celui de toutes les autres bases de données, d'après la SPEC (2004).

MySQL – la haute fiabilité

La fiabilité et la contribution au fonctionnement régulier du système sont des caractéristiques essentielles que l'on exige d'un logiciel. Cela est d'autant plus vrai dans le cas d'une base de données. MySQL ne déçoit nullement dans ce domaine et offre de nombreuses fonctions qui contribuent à assurer le fonctionnement continu de la base de données.

Une installation MySQL autonome présente une fiabilité élevée, de nombreuses installations MySQL n'ont présenté qu'un temps d'interruption très faible, voire nul, sur plusieurs années. Comme nous l'avons vu précédemment, beaucoup de grandes entreprises utilisent la réplication de MySQL pour mettre en place des architectures en mode scale-out capables d'assurer des niveaux très élevés de redondance, offrant ainsi un fonctionnement sans aucune interruption pour des dizaines, voire des centaines de serveurs.

Le serveur MySQL lui-même est un logiciel quasiment exempt de bogues, ce qui contribue énormément à la fiabilité globale. La qualité de MySQL a été mesurée par des organismes de tests indépendants, notamment KlocWork, Coverity et Reasoning. Ces tests ont montré que, grâce à la haute qualité du processus de développement open source, MySQL ne présentait que 1/6^e du taux d'erreurs de logiciels propriétaires comparables. MySQL certifie ses logiciels auprès d'installations de tests externes. Par ailleurs, étant open source, il bénéficie d'une communauté d'utilisateurs considérables, avec plus de 10 millions d'installations qui testent, valident et améliorent constamment le logiciel. Pour ne donner qu'un seul exemple, le bêta test effectué par Microsoft en 2005 pour son produit SQL Server 2005 comptabilisait 300 000 téléchargements de son logiciel, tandis que le bêta test simultané portant sur MySQL 5.0 représentait 2 millions de téléchargements.

Pour tout ceux qui ont besoin d'une disponibilité d'utilisation la plus élevée possible, MySQL Cluster permet de mettre en place des solutions de haute disponibilité à « cinq 9 » (99,999% de disponibilité). Avec son architecture shared-nothing, MySQL Cluster est une base de données résidant en mémoire qu'utilisent ceux qui ne peuvent absolument pas se permettre de voir leur système tomber en panne. MySQL Cluster est particulièrement bien adapté à des applications de télécommunications et aux systèmes de commerce électronique transactionnel.

Un logiciel simple mais riche en fonctions

L'une des priorités de MySQL est d'assurer la convivialité et la simplicité d'utilisation de son logiciel. Quelle que soit la plate-forme supportée, un professionnel des bases de données découvrant MySQL sera en mesure de télécharger et d'installer complètement le logiciel en moins de 15 minutes. Une fois celui-ci installé, un administrateur ou un développeur de bases de données notera que le respect des normes ANSI SQL lui permet d'être immédiatement productif. Les utilisateurs de MySQL apprécieront également l'existence d'un éventail d'outils graphiques d'administration et de développement fonctionnant sous Windows, Linux et Mac, qui améliorera leur productivité et simplifiera l'interaction avec la base de données.

Bien que MySQL reste simple en termes d'installation, de configuration et de gestion, cela ne signifie pas pour autant qu'il s'agisse d'un SGBD pauvre en fonctionnalités. Bien au contraire, MySQL n'a rien à envier à l'immense majorité des bases de données propriétaires lorsqu'il s'agit de fonctionnalités les plus courantes. Un rapide coup d'œil à quelques-unes des fonctions de MySQL suffit pour s'en convaincre.

Fonctions de bases de données	Possible avec MySQL
Support multi plates-formes (Linux, Unix, Windows, Mac ; 32 & 64 bits)	✓
Redémarrage/récupération après incident automatiques	✓
Gestion automatique du stockage (auto-expansion, gestion des annulations)	✓
Prise en charge sérieuse de la sécurité (SSL, privilèges objets à granularité fine)	✓
Chiffrement/déchiffrement des données intégré	✓
Sauvegarde en ligne avec récupération point-in-time	✓
Prise en charge de la base de données en cluster	✓
Capacités de réplication	✓
Moteurs de stockage additionnels	✓
Prise en charge d'une gamme complète de types de données (caractère variable, BLOB, horodatage, etc.)	✓
Support géospatial	✓
Tables Archive et comprimées	✓
Prise en charge d'Unicode et d'UTF-8	✓
Puissantes fonctions d'indexation (en cluster, b-tree, hachage, texte intégral)	✓
ANSI SQL, curseurs, joins, sous-requêtes	✓
Procédures stockées, déclencheurs, fonctions SQL et définies par l'utilisateur	✓
Vues actualisables	✓
Transactions au niveau ACID	✓
Verrouillage illimité au niveau des lignes	✓
Lecture instantanée/constante (les lecteurs ne bloquent pas les scripteurs ou vice versa)	✓
Transactions distribuées	✓
Intégrité référentielle assurée par le serveur	✓
Types de données numériques haute précision	✓
Support VLDB (au niveau du téraoctet)	✓
Utilitaire de chargement de données haut débit	✓
Caches mémoire multiples	✓
Optimiseur de coûts	✓
Dictionnaire de données	✓
Prise en charge des environnements intégrés et profondément intégrés	✓

Suggestions pratiques pour une migration facile vers MySQL

Démarrer un nouveau projet de développement sur la plate-forme MySQL se fait sans aucune difficulté, mais qu'en est-il de la migration d'une base de données existante vers MySQL ? Quel programme une entreprise doit-elle suivre pour être certaine que la transition sera aussi fluide que possible ?

Nombre d'entre elles souhaiteront pouvoir utiliser MySQL en conjonction avec leur base de données propriétaire existante. C'est par exemple le cas de Sabre et de CitySearch, qui ont utilisé MySQL pour créer une solution de bases de données à plusieurs niveaux qui rend plus économique le déploiement de nouvelles applications en mode scale-out sans avoir à sacrifier les investissements déjà réalisés pour leur système de base de données existant. Bien qu'il ne s'agisse pas là, à proprement parler, d'un projet de migration, c'est souvent une approche pragmatique permettant de maîtriser les coûts toujours croissants de licence et de maintenance de bases de données propriétaires.

Le tableau ci-dessous illustre une stratégie de migration pas à pas pouvant être utilisée par quasiment tout le monde pour migrer en toute confiance une base de données existante vers MySQL. Pour chaque étape, une liste des tâches décrit la série d'activités à réaliser (et comment mener à bien chacune de ces étapes).

1	2	3	4	5
Documenter les sources	Concevoir les cibles	Concevoir les flux ETL	Construire/effectuer les migrations	Tests de performances
<ul style="list-style-type: none"> • Comptes utilisateurs • Entités • Relations • Types de données • Règles (données, etc.) 	<ul style="list-style-type: none"> • Comptes utilisateurs • Entités • Relations • Types de données • Règles (données, etc.) 	<ul style="list-style-type: none"> • Mouvements simples • Transformations de données • Agrégations • Profilage/nettoyage des données • Faire migrer le code ? 	<ul style="list-style-type: none"> • Opérations d'ETL • Planifiées/Récurrentes • Validation 	<ul style="list-style-type: none"> • Simulation • Analyse des chiffres • Vérification
<ul style="list-style-type: none"> • Scripts/manuels • Modèles de données • Modèles, rapports, etc. 	<ul style="list-style-type: none"> • Scripts/manuels • Modèles de données • Modèles, rapports, etc. 	<ul style="list-style-type: none"> • Scripts/Fournis par l'éditeur • Partenaire tactique • Dans toute l'entreprise 	<ul style="list-style-type: none"> • Scripts/manuels • Partenaire tactique • Dans toute l'entreprise 	<ul style="list-style-type: none"> • Manuel • Partenaire tactique • Automatisés

Figure 3 - Une stratégie de migration pas à pas pour migrer une base de données existante vers MySQL.

Selon la complexité de l'application, certaines étapes seront plus délicates que d'autres. Par exemple, la toute première étape – Documenter les sources – sera très longue à effectuer dans le cas de bases de données contenant des centaines ou des milliers d'objets de données (tables/index).

Pour aider les professionnels de MySQL à migrer facilement des bases de données de toutes tailles et de tous niveaux de complexité, MySQL propose un puissant outil, nommé Migration Toolkit, qui réduit considérablement le temps et le travail exigés par un projet de migration. Parcourons rapidement chaque étape du cycle pour comprendre comment l'outil Migration Toolkit de MySQL peut largement accélérer le processus

de migration. Le Migration Toolkit est totalement extensible et prend en charge la migration vers MySQL 5.0 de bases de données Oracle, SQL Server, Access, ou de toute autre base de données avec connexion JDBC.

MySQL offre également une solution de conseil packagée de 5 jours, « Démarrage d'une migration vers MySQL », destinée à tous ceux qui souhaitent qu'un conseiller MySQL certifié les aide à analyser leur infrastructure actuelle, à identifier les éventuels problèmes de migration, à migrer un échantillon de leurs données et à effectuer des tests de performance et de charge à partir d'exemples de requêtes. Vous trouverez davantage d'informations à <http://www.mysql.fr/why-mysql/migration/>

Étape 1 – Documenter les sources

Documenter une structure de base de données existante peut être délicat si l'on tente l'opération à l'aide de procédures manuelles standard. Bien que toutes les bonnes bases de données disposent d'un dictionnaire de métadonnées, la procédure consistant à extraire manuellement toutes ces métadonnées (tables, colonnes, index, etc., descriptions) peut prendre un temps considérable.

La meilleure méthode est d'utiliser une procédure informatisée de rétro-ingénierie qui dresse automatiquement le catalogue de toutes les métadonnées nécessaires à la conversion. On peut avoir recours pour cela à un bon outil de modélisation des données. Un certain nombre d'éditeurs en proposent, par exemple Sybase/Quest avec Powerdesigner et Embarcadero avec ER/Studio, lesquels peuvent effectuer la rétro-ingénierie de multiples sources de données telles qu'Oracle, SQL Server, DB2 ou autres.

L'outil graphique Migration Toolkit de MySQL dispose d'un composant de rétro-ingénierie absolument identique à celui des meilleurs outils de modélisation des données. L'utilisateur commence par fournir des informations de connexion au système de base de données source depuis une interface conviviale, puis se connecte et réalise la rétro-ingénierie des schémas destinés à la migration.



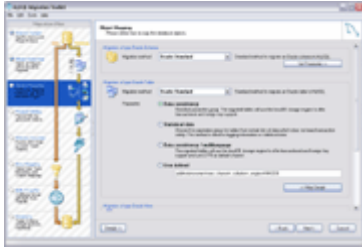
Avec MySQL Migration Toolkit, obtenir les métadonnées nécessaires à partir de la base de données source est un jeu d'enfant.

Étape 2 – Concevoir les cibles

Une fois les métadonnées de la source obtenues, l'étape suivante consiste à concevoir la base de données MySQL cible. Il s'agit essentiellement de traduire les objets sources et leurs propriétés (telles que les types de données de colonnes) en compléments MySQL. Comme vous pouvez l'imaginer, cette étape peut s'avérer extrêmement longue et périlleuse si l'on tente de l'effectuer manuellement, la plupart des bases de données ayant des milliers de propriétés d'objets à convertir.

Heureusement, le MySQL Migration Toolkit convertit automatiquement toute base de données source ayant subi une opération de rétro-ingénierie en son homologue MySQL, avec toutes les traductions de types de données et les conversions nécessaires. L'utilisateur conserve la maîtrise absolue des objets à migrer et

détermine lui-même les spécifications MySQL à utiliser, par exemple en choisissant le moteur de stockage sous-jacent.



Comme nous l'avons déjà vu, cette étape peut se révéler extrêmement délicate. Si elle n'a pas été effectuée correctement, un DBA pourra se trouver contraint de faire marche arrière pour concevoir à nouveau les objets ou réparer les échecs de transfert de données dus à un mappage de type de données invalide. Toutefois, si l'on utilise le MySQL Migration Toolkit, ces problèmes ne seront plus qu'un mauvais souvenir, l'outil étant capable d'effectuer correctement le travail dès la première fois.

Étape 3 – Concevoir les flux ETL

Une fois que les métadonnées sources ont été comprises et la base de données MySQL cible conçue, il va falloir, à l'étape suivante, établir le processus de migration des données elles-mêmes. La phase ETL (Extract, Transform & Load, *extraire, transformer et charger*) peut s'avérer assez complexe selon ce que l'on souhaite accomplir. Outre le MySQL Migration Toolkit, on trouve sur le marché de nombreux outils d'ETL puissants qui permettent avec une grande souplesse de décider comment déplacer, agréger, mapper et transformer les données depuis les bases de données sources vers les bases cibles. Pour tous ceux qui procèdent en permanence à des migrations de bases de données (par exemple pour transférer des données depuis les systèmes source vers les entrepôts de données), l'achat de ces outils représentera un bon investissement.

S'il ne s'agit que d'un transfert unique, cependant, le prix de ces outils d'ETL sera dissuasif. Heureusement, le MySQL Migration Toolkit pourra encore une fois vous venir en aide. Il dispose de nombreuses options permettant de déplacer une base de données source vers MySQL et son prix est imbattable puisque le Migration Toolkit est édité sous la licence open source GPL.

Une des stratégies de migration des données consiste à faire migrer tous les objets et données existants depuis la base de données source vers une base de données MySQL relais. Une fois ceux-ci en sécurité sur le serveur de la base de données MySQL, un DBA ou un développeur pourra créer des procédures stockées ou tout autre code de migration capable de déplacer et de manipuler les données de façon sélective depuis la base de données relais vers une autre base de données MySQL qui servira lors de phases ultérieures de développement ou de production.

Une autre option consiste à utiliser les possibilités de script du MySQL Migration Toolkit pour traiter les données lors de leur transfert depuis la base de données source vers la base de données MySQL cible. Bien que la plupart des utilisateurs décident d'employer l'interface graphique du Migration Toolkit pour faire migrer leurs bases de données en mode pointer/cliquer, les utilisateurs chevronnés pourront exploiter l'architecture sous-jacente pour écrire des scripts décrivant des procédures ETL plus complexes.

Étape 4 – Construire/effectuer des migrations

L'étape suivante de la procédure de migration est le transfert proprement dit des objets de données et, le cas échéant, des données sous-jacentes. De nombreux facteurs rentrent ici en ligne de compte, par exemple le volume des données transférées, l'ordre dans lequel les objets et leurs données associées sont transférés, etc.

Le MySQL Migration Toolkit offre une interface extrêmement conviviale pour réaliser la migration de la base de données elle-même, tout en garantissant que tous les mouvements de données seront effectués dans le bon ordre et de la façon la plus efficace possible. Pour des mouvements de données de grande ampleur, le Migration Toolkit déplace les données en les regroupant et vérifie que toutes les propriétés de la base et des objets contribuant à l'obtention d'un temps de chargement rapide sont correctement paramétrées.

Les utilisateurs du Migration Toolkit sont tenus informés de la progression de la migration à l'aide de barres de progression faciles à lire et d'autres messages d'information. Une fois l'opération achevée, l'utilisateur voit s'afficher un résumé de toutes les tâches effectuées.

Étape 5 – Tests de performance

Une fois que tout a été extrait de la base de données source vers MySQL, il est capital de procéder à une série de tests afin de vérifier que les performances sont acceptables. Cette étape est peut-être la plus négligée de toutes celles du cycle de migration. Elle n'était jusqu'ici pas facile à accomplir.

Des tests de performance correctement effectués détectent les problèmes de performances que des tests utilisateur et d'assurance-qualité inappropriés ne perçoivent pas. En résumé, ces tests simulent ce qui se passera concrètement en grandeur nature. Ils pèsent sur la base de données d'une façon qui ne pourrait être obtenue qu'en ouvrant grandes les vannes de la communauté des utilisateurs professionnels.

Des tests de performances intelligents utilisent les composants suivants pour effectuer une simulation réaliste de ce que subira une base de données lors de sa future utilisation en production :

- **Présence utilisateurs attendue** – il est essentiel que le test simule le nombre de connexions utilisateurs attendues lors des périodes de pointe et durant les heures normales d'ouverture. Il s'agit du principal point sur lequel échouent les méthodes manuelles consistant à choisir une sous-catégorie d'utilisateurs pour tester une base de données et une application. La base de données pourra fonctionner de façon satisfaisante avec une dizaine de connexions utilisateurs mais s'effondrer si 400 personnes se connectent au système.
- **Activité utilisateurs répétée** – une fois que le nombre prévu de sessions utilisateurs se sont connectées à la base de données, encore faut-il que ces sessions « fassent quelque chose » pour que le système subisse une pression. Et ce « quelque chose » soit accompli plus d'une fois. Les sessions connectées doivent, en totalité ou en partie, effectuer des tâches de façon répétitive comme cela se passerait lors d'une journée de travail normale. Pour un système OLTP, il peut s'agir de la saisie de plusieurs commandes. Pour un entrepôt de données, cela peut être la production de requêtes analytiques de longue haleine. Le tout est que le travail soit répétitif, de façon à ce que la base de données subisse des assauts répétés.
- **Période prolongée** - une fois que vous disposez d'un nombre déterminé de sessions accomplissant un travail répétitif, vous devez vous assurer que ce travail se poursuit sur une durée significative pour que le test soit valide. Concrètement, vous cherchez à mettre à jour des problèmes qui n'apparaissent qu'au bout d'un certain

temps. Par exemple, une table SQL Server peut ne pas se fragmenter à l'issue de 30 minutes de travail OLTP, mais se fragmenter soudain de façon alarmante après 2 heures ou plus d'action répétée.

- Volume attendu de données de production – pour que le test soit réellement validé, vous devez charger votre base de données de données de test d'une taille comparable à celle qu'elle atteindra réellement une fois en production. Ce composant est facilement réalisable si une base de données de production existante a migré vers MySQL.

La plupart des DBA s'interrogent fortement sur la façon pratique de mettre en place la totalité des points ci-dessus. Il existe bien des outils d'évaluation open source, mais la plupart d'entre eux présentent des défauts qui les rendent inadaptés à des tests de performance précis.

Des outils proposés par des éditeurs tiers peuvent aider, certains plus efficacement que d'autres. Embarcadero Technologies édite Extreme Test, un outil de tests de performance extrêmement souple qui peut effectuer des tests de charge sur la plupart des bases de données, y compris MySQL, ainsi que sur tous les systèmes d'exploitation les plus répandus. Dans le haut de gamme, on trouve les outils de tests de charge édités par Mercury Interactive.

Quest Software commercialise un excellent outil de tests de performance, Benchmark Factory. D'une grande simplicité d'utilisation, celui-ci dispose d'assistants en mode pointer/cliquer qui guident étroitement l'utilisateur tout au long de la procédure de configuration et de réalisation d'un grand nombre de tests TPC standard pour MySQL, ainsi que d'autres tests de performance évolutifs. Une fois les tests achevés, Benchmark Factory offre des utilitaires d'analyse qui permettent de comparer facilement les résultats des tests et de les exporter dans Microsoft Excel.

Conclusion

D'un point de vue commercial comme d'un point de vue technique, il est facile de répondre à la question de savoir s'il est intéressant de migrer vers MySQL. D'innombrables entreprises modernes réalisent d'énormes économies à l'aide de ce logiciel tout en faisant fonctionner avec une grande facilité leurs systèmes de bases de données les plus exigeants grâce à son architecture en scale-out.

Bien qu'une migration de base de données soit rarement facile, suivre le cycle de migration des données en cinq étapes décrit dans le présent document vous aidera à assurer une transition fluide et, au final, une migration réussie. En outre, MySQL offre pour cela son Migration Toolkit en téléchargement gratuit, lequel réduit considérablement le temps nécessaire pour effectuer une telle migration, ainsi que la quantité d'erreurs qui résultent normalement d'opérations aussi complexes lorsqu'elles sont tentées manuellement. Enfin, à l'issue de la migration, un DBA ou un responsable projet devra effectuer des tests de performances adéquats afin de s'assurer que les niveaux de performance seront suffisants pour faire face à la charge que subira le système lors de sa mise en production.

Ressources

Aides MySQL à la migration

MySQL Migration Toolkit

<http://www.mysql.fr/products/tools/migration-toolkit/>

Faites migrer graphiquement une base de données Oracle, SQL Server, Microsoft Access ou une autre base accessible par JDBC vers MySQL.

MySQL Migration Central

<http://www.mysql.fr/products/tools/migration-toolkit/>

Apprenez comment migrer facilement et rapidement votre base de données existante vers MySQL.

Forums migration

<http://forums.mysql.com>

Échangez des informations avec d'autres professionnels MySQL concernés par la migration de leur base de données actuelle.

Études de cas

Sagem Monetel démocratise la mise à jour des terminaux électroniques avec MySQL

http://www.mysql.fr/why-mysql/case-studies/sagem_fr.php

Ce spécialiste internationalement reconnu élabore et commercialise des terminaux de transaction électronique. Pour gérer à distance la mise à jour logicielle des terminaux électroniques, Sagem Monetel propose à ses clients son logiciel TMS. Précédemment stockées dans une base de données traditionnelle du marché, les données résident désormais dans MySQL 5.0

Utel traite 10,000 requêtes/seconde grâce à un déploiement horizontal de MySQL Network

http://www.mysql.fr/why-mysql/case-studies/utel_fr.php

Le pari d'Utel: développer la plus grande communauté de chatteurs sur téléphone mobile de France! Entièrement basé sur MySQL, le système traite en moyenne 3,000 requêtes par seconde avec des pics à 10,000 requêtes par seconde. Soit plus de 300 millions de requêtes par jour.

Livres blancs

Guide sur la Diminution du Coût Total de Possession des Bases de Données

<http://www.mysql.fr/why-mysql/white-papers/tco.php>

Il y a quelque temps Computerworld avait révélé comment MySQL était devenue la base de données open source la plus populaire au monde, et pourquoi les sociétés décidées à réduire leurs coûts d'opérations l'utilisent.

MySQL permet de réduire le Coût Total de Possession (TCO) du logiciel de base de données en diminuant :

- Les coûts de licence de base de données de plus de 90%
- Le temps d'arrêt de 60%
- Les dépenses sur le matériel de 70%
- Les coûts d'administration, d'ingénierie et de support jusqu'à 50%

Au coeur de MySQL 5.0, MySQL 5.0 vu par les DBA

http://www.mysql.fr/why-mysql/white-papers/mysql_wp_inside50.php

MySQL 5.0 représente un véritable bond en avant pour le système de gestion de bases de données open source le plus populaire au monde. On sait que MySQL est depuis des années la base de données de prédilection lorsqu'il s'agit de gérer des sites Web à fort volume et des applications de bases de données embarquées. Mais la version 5.0, qui introduit de nouvelles fonctionnalités tout à fait exceptionnelles, va encore accélérer son adoption par les entreprises. Les progrès qu'elle apporte dans le domaine du développement applicatif, du traitement des transactions, de l'intégrité des données et de la facilité d'administration placent le serveur de base de données MySQL sur un pied d'égalité avec des bases de données propriétaires coûtant de nombreuses fois son prix.

Ce document commence par une présentation technique de MySQL du point de vue des DBA (Data Base Administrators, administrateurs de bases de données), puis expose les nouvelles fonctionnalités de la version 5.0. Il montre que les améliorations apportées à cette nouvelle version à partir d'un moteur de base de données déjà primé aidant les entreprises à mieux gérer les systèmes transactionnels exigeants, les applications d'entreposage de données de grande taille et les sites Web à fort trafic.

Rapports d'analystes

Your Open Source Strategy (Votre stratégie open source), Forrester : Schadler, septembre 2003

<http://www.mysql.com/it-resources/analyst-reports/forrester-open-source-databases.php>

Les entreprises sont intriguées par les logiciels open source, mais restent bloquées par les mythes circulant sur leur coût, leur support et leurs risques. Les entreprises astucieuses sauront dépasser ces mythes pour se procurer les logiciels ouverts dont elles ont besoin.

So Much Database So Little Money (Une énorme base de données pour très peu d'argent), Hurwitz Associates, avril 2005

<http://www.mysql.com/it-resources/analyst-reports/hurwitz-vantage-point.php>

Y a-t-il de bonnes raisons de payer des frais de licence vertigineux pour une base de données d'entreprise ? À vrai dire, cela dépend des circonstances, mais dans la plupart des situations, la réponse est « Non ». Y a-t-il de bonnes raisons d'utiliser une base de données totalement gratuite au sein de l'entreprise ? En fait, cela aussi dépend des circonstances. Après tout, de nombreuses entreprises utilisent un navigateur gratuit (Firefox), un serveur Web gratuit (Apache) et un système d'exploitation gratuit (Linux). S'agissant des bases de données,

néanmoins, la situation est en général plus compliquée. Le support technique peut et doit être une préoccupation majeure et un support 24h/24, 7j/7 avec réponse garantie n'est jamais gratuit.

The End of Database Licensing? (*La fin des licences de bases de données ?*) Meta Group, avril 2005

<http://www.mysql.com/it-resources/analyst-reports/meta-group-db-licensing.php>

Alors que la concurrence autour de l'open source se fait plus âpre, c'est le support (par exemple au niveau technique, en termes d'implémentation, auprès des fournisseurs de services Internet) qui déterminera la vitesse à laquelle il sera adopté et qui décidera du vainqueur. MySQL Network est une offre de support innovante qui menace les modèles classiques de licences logicielles du commerce. MySQL Network est une offre qui devrait accroître la vitesse de pénétration de MySQL sur le marché de l'entreprise et marquer le « début de la fin » pour les frais de licences logicielles en tant qu'élément facturé séparément.