



SGBD/R LIBRES

PAR ALAIN BASTIDE

Les bases de données Open Source relèvent tous les défis

Ils ont dû faire leurs preuves pendant des années. Mais les SGBD/R libres sont désormais acceptés en production sur des projets critiques. La concurrence se développe. Les outils se spécialisent. Et les entreprises revendiquent désormais le recours à ces outils Open Source. 2008, année de la maturité ?



L EN AURA FALLU des années... mais les systèmes de gestion de bases de données relationnelles (SGBD/R) Open Source ont fini par s'imposer comme une alternative viable aux outils propriétaires. Malgré leur faible coût de licence, en 2007 ils représentaient près de 6% d'un marché estimé à 15 milliards de dollars par Gartner. Les SGBD/R Open Source ont généré 300 millions de dollars de revenus en 2006, mais ils atteindront un milliard dès 2008 selon Forrester.

Au delà des aspects financiers, c'est la dynamique du marché qui est la plus intéressante. *« Alors que le marché des SGBD/R a progressé dans son ensemble de 8% entre 2006 et 2007, celui des outils Open Source a bondi de 30% sur la même période »*, analyse Éric Soares, directeur général d'Ingres en France.

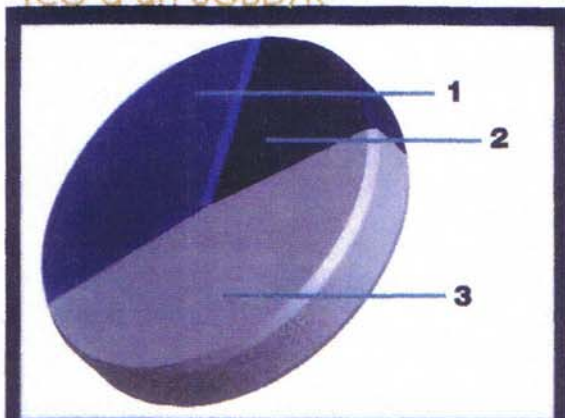
Si l'on compte le nombre de serveurs déployés, la percée de l'Open Source est encore bien plus fulgurante. Icône des SGBD/R libres, MySQL revendique par exemple plus de 11 millions de serveurs déployés. Le taux de pénétration de MySQL était de 40% des entreprises en 2006; 66% déploient ou planifient de déployer MySQL en 2008 estime Gartner. Ce qui fera de MySQL la troisième base de données la plus déployée en entreprise dès cette année.

Les SGBD/R en chiffres

- 90 % des entreprises ont plus de deux SGBD/R différents déployés (Forrester).
- 70 % des entreprises utiliseront un SGBD/R Open Source d'ici 2008 (Gartner).
- 33 % des utilisateurs d'Oracle utilisent aussi MySQL en production - Independent Oracle User Group
- 52 % des grandes entreprises américaines utilisent ou comptent utiliser MySQL (Forrester).
- 35 % des SGBD/R Open Source déployés hébergent des données critiques (Forrester).
- 20 millions de serveurs MySQL déployés et jusqu'à 60 000 téléchargements par jour sur mysql.com (MySQL AB).



TCO d'un SGBD/R



- 1 : 28 % de coût d'indisponibilité,
 2 : 15 % de coût de licence,
 3 : 40 % de coût d'exploitation.

Source : IDC

« Les demandes de formations MySQL ont plus que doublé entre 2006 et 2007 », confirme Cyril Pierre de Geyer, fondateur de l'Observatoire des Logiciels Libres et responsable R&D pour l'organisme de formation Anaska.

Si le succès est au rendez-vous, c'est que les entreprises cherchent toujours à abaisser le coût total de possession (TCO) de leurs données. Selon AMR Research, le TCO des outils propriétaires reste trop élevé. « Le coût d'acquisition représente encore 15% du TCO. C'est le principal critère de choix d'un SGBD libre. Et ce, malgré les ristournes parfois importantes des éditeurs propriétaires », confirme Cyril Pierre de Geyer.

DES OUTILS MATURES

L'absence de coût de licence n'est cependant pas un critère suffisant pour retenir une base de données libre, plutôt qu'un outil propriétaire. Car malgré leur coût élevé, les SGBD/R propriétaires sont, en général, particulièrement efficaces. « Lorsque nous sommes en compétition avec Oracle et Microsoft, le coût est l'attraction initiale pour les clients, mais ils continuent à utiliser notre logiciel pour sa performance et sa fiabilité », illustre Bertrand Mattheï, directeur marketing EMEA de MySQL.

Au-delà du faible coût de licence, c'est donc l'arrivée à maturité des outils Open Source qui explique la croissance de ce marché depuis deux ans. « MySQL couvre 70% des besoins des utilisateurs d'Oracle », constate Yann Neuhaus, expert SGBD/R de Trivadis, certifié Oracle RAC et MySQL 5. Comme MySQL est gratuit, « simple à installer et à administrer, et peut même offrir de meilleures performances », il constitue une alternative très sérieuse à Oracle pour certains types de projets. « Mais pas encore tous », tempère Yann Neuhaus. « Dans des domaines tels que la consolidation et les architectures garantissant la montée en charge, Oracle RAC est très puissant », explique-t-il.

Les jalons de l'histoire des SGBD/R Open Source

2004

Février : MySQL Administrator, première console d'administration graphique pour MySQL.

Avril : MySQL présente la première version de MySQL Cluster, une base de données distribuée qui s'exécute intégralement en mémoire. Un concurrent Open Source d'Oracle RAC.

Septembre : pour répondre à la pression de MySQL, Oracle et IBM proposent des versions limitées de leurs SGBD/R à partir de 4000 euros (10g Standard Edition One et DB2 UDB Express 8.1). Microsoft avait déjà annoncé SQL 2005 Express, une version gratuite et limitée de SQL Serveur quelques mois auparavant.

2005

Janvier : PostgreSQL 8.0 fonctionne nativement sur Windows.

Avril : MySQL investit le marché du décisionnel avec MySQL Migration Suite, un outil très proche d'un ETL et de nombreux moteurs OLAP fournis par des partenaires.

Octobre : MySQL 5.0 ajoute procédure stockée, vue et triggers.

Octobre : Oracle acquiert Innobase qui développent InnoDB, principal moteur transactionnel de MySQL.

Novembre : PostgreSQL 8.1 vise les entrepôts de données.

2006

Janvier : rentable depuis six mois, MySQL annonce 2 000 nouveaux clients depuis 2005.

Janvier : IBM présente DB2 Express-C, une version gratuite et limitée de DB2.

Mars : MySQL rejoint la fondation Eclipse.

Avril : MySQL embauche Jim Starkey, le père d'Interbase (Firebird), pour créer Falcon son nouveau moteur transactionnel - afin de proposer une alternative à InnoDB. Création de MySQL Forge.

Mai : MySQL ouvre officiellement son API pour permettre à ses partenaires de développer des moteurs de stockage tiers. Moteur MySQL pour IBM DB2 sur iOS/5. MySQL Cluster atteint 100 000 transactions par minute.

Octobre : Plug-in MySQL pour Visual Studio 2005.

2007

Janvier : lancement de MySQL Enterprise Unlimited, les licences de site MySQL Enterprise Unlimited sont désormais disponibles pour le prix d'un seul processeur d'Oracle Enterprise Edition.

Février : MySQL Cluster Carrier Grade Edition vise le marché des télécoms.

Octobre : avec IceBreaker, Ingres propose une appli virtuelle comprenant son SGBD/R, la suite décisionnelle de Jaspersoft et l'ETL de Talend.

Octobre : MySQL Enterprise « Automne 2007 » permet de superviser et d'automatiser les opérations de réplication.

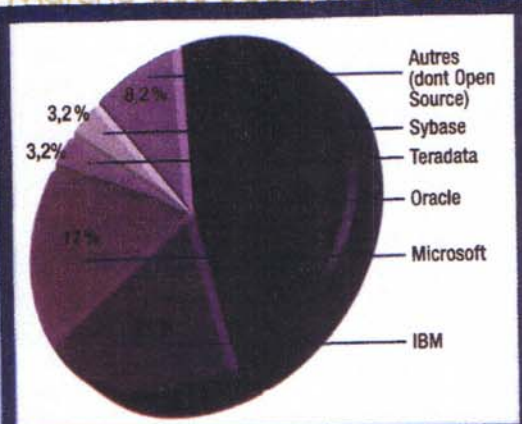
Novembre : Google va contribuer au développement de MySQL 6.0.

Décembre : EnterpriseDB Advanced Server 8.3 (basé sur PostgreSQL) améliore encore sa compatibilité avec Oracle avec deux nouveaux émulateurs pour SQL*Plus et SQL*Loader.





Marché des SGBD/R en 2006



En 2006, le chiffre d'affaires cumulé (licence, support et services) de tous les éditeurs représente 15,2 milliards de dollars.

Source : Gartner 2007

Pour Peter Kastner, vice-président du cabinet d'analyse Aberdeen Group, « les bases de données à code source libre se sont radicalement améliorées tant sur les plans de l'évolutivité, de la robustesse et de l'administrabilité. Elles méritent d'être comparées aux bases de données d'entreprise », estime-t-il. L'âge de la maturité, donc.

Cependant, conscient qu'il ne détrônara pas les outils propriétaires du jour au lendemain, MySQL ne se positionne pas en frontal d'Oracle et d'IBM. « Nos clients utilisent souvent notre moteur comme un complément à leur base propriétaire existante [NDLR : 1/3 des utilisateurs Oracle ont déployé MySQL], pour de nouveaux projets qui nécessitent une architecture de données moderne », explique Bertrand Matthélié.

MYSQL N'EST PLUS LE SEUL MAÎTRE À BORD

MySQL se spécialise peu à peu – Web 2.0, SaaS, équipementiers et opérateurs télécoms, décisionnel, etc. – car, même s'il reste

« MySQL est le seul SGBD/R conçu spécifiquement pour les besoins actuels des entreprises »

Plusieurs évolutions convergent et vont toucher le marché des SGBD/R en 2008 : une croissance exponentielle du nombre d'utilisateurs d'Internet (1 milliard) et de téléphones mobiles (3 milliards), une proportion toujours plus importante d'applications d'entreprise en ligne avec l'évolution du Web 2.0 vers l'entreprise 2.0, et enfin la croissance rapide du marché SaaS (Software as a Service). Ces trois tendances nécessitent un nouveau type de SGBD/R, en ligne, qui doit fournir un accès instantané, une très faible latence et une haute performance. Il doit être simple, rapide et économique à mettre en place pour tester différentes nouvelles offres et services, mais doit également pouvoir monter en charge très vite pour gérer les demandes massives générées par le succès de certaines de ces offres. MySQL est la seule base de données qui a été conçue spécifiquement pour répondre aux besoins du monde moderne en ligne.

Bertrand Matthélié, Sr. Director of Marketing, EMEA



Les principaux SGBD/R Open Source pour l'entreprise

Il existe aussi...

SGBD généralistes : Firebird (communauté / Borland), MaxDB (SAP).
Produits embarqués : SQLite, BerkeleyDB (SleepyCat / Oracle), MySQL Embedded (MySQL AB), Derby (Apache / IBM).
Informatique distribuée : MySQL Cluster (MySQL AB) et SolidDB (IBM).

le leader incontesté, la concurrence Open Source se développe. Un autre signe de maturité du marché.

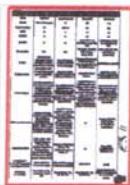
Depuis deux ou trois ans, de nouveaux éditeurs ont vu le jour. C'est notamment le cas d'EnterpriseDB, qui « professionnalise » PostgreSQL, et d'Ingres, qui revient sur le devant de la scène. EnterpriseDB se positionne comme l'alternative Open Source à Oracle. L'éditeur promet 100% de compatibilité, des performances identiques, et un outillage adapté aux besoins des entreprises à une fraction du coût des outils d'Oracle. Cette stratégie de « me-to » semble porter ses fruits. EnterpriseDB a connu une croissance de 250% de ses ventes de supports entre 2006 et 2007 auprès de clients prestigieux comme Sony.

Revenant sur le devant de la scène, Ingres crée de son côté des appliances virtuelles pour adresser des besoins précis. Son appliance IceBreaker intègre par exemple le SGBD/R Ingres 2006, la suite décisionnelle JasperBI de Jaspersoft, et l'ETL de Talend au-dessus d'une distribution Linux rPath. Ingres a ainsi conquis des entreprises comme Air France, Renault et Total.

Enfin, l'utilisation de SQLite dans Google Gears, AIR d'Adobe et Firefox 3.0 a aussi ravivé l'intérêt pour cette base de données embarquée.

Ces nouveaux concurrents commencent à faire peur aux éditeurs traditionnels. Oracle a donc racheté InnoDB et SleepyCat. « Il est possible qu'Oracle ait souhaité freiner MySQL dans sa croissance au moment même où MySQL et SAP tentaient un rapprochement technologique », rappelle Stéphane Schildknecht, président de PostgreSQL.fr. Privé de ses moteurs transactionnels, MySQL redevient une base de données par fichiers indexés. Heureusement, il existe de nombreuses alternatives à InnoDB – que la licence GPL protège par ailleurs d'une action trop agressive de la part d'Oracle. SleepyCat est de son côté une entreprise très rentable puisqu'elle commercialise BerkeleyDB, la base de données embarquée la plus utilisée dans le monde. Plus récemment, c'est IBM qui a racheté Solid Information Technology pour compléter DB2.





| | MySQL | Ingres | EnterpriseDB | PostgreSQL |
|--|--|--|--|---|
| Éditeur | MySQL AB | Ingres Corporation | EnterpriseDB | Communauté |
| Licence | GPL et commerciale | GPL | GPL | BSD |
| Dernière version | 5.1 | 2006 R3 | 8.3 | 8.3 |
| SQL92 | oui | oui | oui | oui |
| SQL2000 | nc | oui | oui | oui |
| SQL2003 | nc | oui | >150 critères sur les 164 du support SQL FULL | >150 critères sur les 164 du support SQL FULL |
| OS supportés | tous | tous | Unix, Linux, Mac OSX, xBSD, Windows (depuis 2005) | Unix, Linux, Mac OSX, xBSD, Windows (depuis 2005) |
| Éditions | Enterprise Server (généraliste), Cluster (distribuée) et Cluster Carrier Grade, Embedded (embarquée et OEM) | Edition Community et Edition Enterprise (iso fonctionnelles) | EnterpriseDB Advanced Server Developer, EnterpriseDB Advanced Server Subscription | Pas d'édition particulière |
| Positionnement | nouveaux déploiements pour les applications modernes | alternative à Oracle Database, IBM DB2 UDB et Microsoft SQL Server | alternative à Oracle Database | alternative à Oracle Database, IBM DB2 UDB et Microsoft SQL Server |
| Différentiateurs | architecture modulaire, simplicité, performances, standard « de facto » du marché, gamme complète répondant à des besoins spécifiques, outil de déploiement et de supervision (MySQL Enterprise) | base installée dans les grandes entreprises (20 000 clients), qualité et maturité, approche appliance pour le marché des PME et des serveurs départementaux | support professionnel pour PostgreSQL avec engagement de service, outillage (migration, supervision et développement), performance (+50 % par rapport à la version Open Source selon l'éditeur) et compatibilité Oracle accrue | ouverture (licence BSD n'impose aucune contrainte commerciale), réactivité et qualité du support de la communauté, conformité aux standards SQL |
| Références françaises | Lafarge, Fnac Billelet, Franprix & Leader Price, Skyrock/Skyblog, Europages, Axway, Business Objects, Sagem Monétel, Alcatel-Lucent, Crédit Mutuel, ministère des Affaires étrangères | Inpi, Météo-France, ministère de la Défense, France 3, Eutelsat, France Telecom, Natixis, Société Générale, BNP Paribas, Renault, ST Microelectronics, L'Oréal, Total, SNCF, Air France, Airbus, Eiffage, Leroy Merlin, C.E.A, Areva | nc | Overblog, TDF, ONF, RATP, ministère de l'Équipement, Socopa-Valtero |
| Autres références | Yahoo!, Google, Nasa, Sabre, Booking.com, Wikipedia, Police Nationale de Suède ou The Phone House, RightNow Technologies, F-Secure, Zimbra, Nokia Siemens Networks, Adobe | University of Washington, Barclays, TNT, Rolls Royce, Carnegie Mellon, 3M | nc | Skype, Afilias, myyearbook.com, classmates.com |
| Coût (HT) de la souscription annuelle fournie par l'éditeur | 479 € à 3999 € / serveur selon options | 6218 € / CPU physiques installés dans la machine | 5995 \$ / CPU | gratuit |
| Concurrents propriétaires | Oracle Database 11g, Microsoft SQL Server 2005 | Oracle Database 11g, IBM DB2 UDB | Oracle Database 11g, IBM DB2 UDB | Oracle Database 11g, IBM DB2 UDB, Microsoft SQL Server 2005, Informix |



« Chaque SGBD/R libre possède de vrais différentiateurs »

MySQL et PostgreSQL se détachent clairement du lot par rapport aux autres outils Open Source. Mais ces outils ne se ressemblent pas. Bien au contraire. Chacun possède des caractéristiques techniques vraiment différenciatrices. PostgreSQL dispose d'une architecture organisée en plusieurs processus « parents » et « enfants », ainsi que de très nombreux réglages notamment au niveau de la gestion mémoire. Plus simple, MySQL dispose d'un moteur de réplication natif, dont les performances sont excellentes. Et en plus de son fonctionnement très performants sur Linux, MySQL peut être utilisé en production sous Windows alors que PostgreSQL n'est compatible avec l'OS de Microsoft que depuis janvier 2005 (version 8.0). Cet exemple montre qu'il ne faut pas chercher à comparer les SGBD/R Open Source entre eux, mais plutôt chercher à comprendre à quels usages ils répondent bien pour remplacer un SGBD/R propriétaire plus coûteux – et rarement moins performant.

Philippe Orsier, chef de projet, spécialiste LAMP chez Uperto, filiale Open Source de Devoteam



ABAISSER LE TCO EN SPÉCIALISANT LES OUTILS

« Libres ou propriétaires, les éditeurs facturent tous un support technique pour financer le développement de leurs outils. Les SGBD/R Open Source doivent donc être au niveau des outils propriétaires s'ils veulent conquérir une part du marché », explique Cyril Pierre de Geyer. Après avoir fait leurs preuves sur un plan purement technique, les éditeurs de SGBD/R Open Source ont outillé leurs bases avec des consoles de supervision et des outils d'administration pour alléger la charge des DBA.

Selon IDC, les coûts d'exploitation représentent en effet 40% du TCO d'un SGBD/R (15% pour la licence et 28% pour l'indisponibilité). MySQL propose donc MySQL Enterprise Monitor et EnterpriseDB, DBA Management Server. Les deux éditeurs fournissent également des éditeurs SQL, des outils de migration et de réplication de données. Pour abaisser son coût d'exploitation, MySQL se distingue aussi par sa simplicité. Un point important, car Forrester Research note dans une étude récente que la majorité des entreprises n'utilisent que 35% des fonctions des SGBD propriétaires.

Pour abaisser encore un peu les coûts d'exploitation sur les gros déploiements, les éditeurs et les communautés spécialisent également leurs outils afin d'adresser des segments de marché précis. MySQL est le cas le plus emblématique avec son SGBD/R généraliste MySQL Enterprise Server, sa base de données distribuées haute disponibilité MySQL Cluster, son édition telcos MySQL Cluster Carrier Grade Edition, et MySQL Embedded (embarqué). L'éditeur a également ouvert ses API en mai 2006 pour permettre à

des partenaires de développer des moteurs de stockage spécifiques. Il en existe aujourd'hui une vingtaine, allant de DB2 sous iOS/5, au système de fichier hébergé S3 d'Amazon, en passant par la compression des données d'un datawarehouse (BrightHouse) ou la fédération de serveurs distribués (Federated).

APPLICATIONS CRITIQUES ET SAAS AUSSI

Au final, les différents SGBD/R Open Source du marché répondent à quasiment tous les besoins des entreprises. Même les plus critiques. De Google à Yahoo! en passant par YouTube, Skyblog, Facebook, et Wikipedia, MySQL « motorise » les plus gros sites Web du monde. Tout comme PostgreSQL, qui équipe Skype, myyearbook.com, classmates.com, overblog.com et bien d'autres.

Depuis deux ans, tous les éditeurs Open Source affichent des prétentions dans le domaine des applications critiques à très forte volumétrie et du décisionnel. « Ingres est traditionnellement utilisé pour les applications transactionnelles métier critiques. C'est aussi une technologie adaptée aux besoins des grands entrepôts de données, en témoigne la solution de Datallegro qui intègre Ingres comme moteur de bases de données », illustre Eric Soares chez Ingres.

Après avoir conquis les entreprises, les équipementiers et opérateurs télécoms (Alcatel-Lucent, Nokia/Siemens, etc.), les acteurs du Web 2.0, et les éditeurs de logiciels (Sage, Axway, Adobe), MySQL vise désormais les fournisseurs d'applications hébergées

« MySQL est simple à utiliser »

Parmi les SGBD/R Open Source, MySQL se différencie des outils propriétaires et de ses concurrents Open Source par sa simplicité d'utilisation et d'installation. La mise en œuvre d'un cluster MySQL est par exemple plus simple qu'avec Oracle RAC. Unique, l'architecture modulaire des moteurs de stockage de MySQL garantit d'excellentes performances. MyISAM est de loin le moteur le plus rapide en lecture intensive et pour le chargement de données décisionnelles, car il ne prend pas en charge les transactions. Or, cette approche est acceptable dans 70% des cas. Et MySQL possède plusieurs moteurs transactionnels parfaitement ACID. Yann Neuhaus – expert en base de données (certifié Oracle RAC et MySQL 5.0) chez Trivadis





« Nos appliances virtuelles accélèrent le déploiement »

Un SGBD/R Open Source évolue en fonction des besoins réels de ses utilisateurs, tandis qu'un produit propriétaire se doit d'être avant tout... différent ! Car il s'agit de « vendre » de nouvelles fonctionnalités. Les outils Open Source sont plus pensés comme des solutions à un problème que comme des produits. Au final, l'éditeur est rémunéré pour la qualité de son accompagnement, sa réactivité et sa capacité à fournir une solution globale. Comme le cycle de vie des outils Open Source est bien plus rapide, il est possible d'intégrer plusieurs produits pour constituer des solutions complètes et performantes en quelques mois. C'est par exemple le cas de nos appliances virtuelles Icebreaker et Ingres BI Appliance conçues avec le concours de RPath et Jaspersoft. Elles réduisent considérablement les délais de mise en œuvre.



Eric Soares, directeur général de Ingres France

(SaaS). « *RightNow, F-Secure, et Zimbra utilisent MySQL pour délivrer des applications critiques à leurs nombreux clients* », explique Bertrand Matthélie, directeur marketing EMEA de MySQL. La plupart du temps MySQL est mis en place pour des nouveaux déploiements plutôt que pour des migrations.

Du côté de PostgreSQL, de grandes entreprises telles que TDF, l'IGN et la RATP commencent à afficher leur retour d'expérience. Un point particulièrement important pour ce SGBD/R 100% communautaire, reconnu pour ses qualités techniques, mais que l'omniprésence de MySQL et le développement commercial d'EnterpriseDB étouffe un peu. « *Il est très important pour nous que les utilisateurs revendiquent leur usage de PostgreSQL, car il ne fait aucun doute que cela pousse d'autres sociétés à franchir le pas* », explique Stéphane Schildknecht.

LE MODÈLE ÉCONOMIQUE SE STANDARDISE

D'Ingres à MySQL en passant par EnterpriseDB, les éditeurs de SGBD/R Open Source ont tous adopté le modèle économique de la souscription. Démocratisé par Red Hat, adopté par MySQL et Zend dans un premier temps, il consiste à fournir un ensemble de produits certifiés et de services associés. Les entreprises n'y recourent généralement pas dans un premier temps. Elles commencent par tester le SGBD/R Open Source sur des projets non critiques. Après une montée en compétence, elles recourent à la souscription pour leurs projets critiques uniquement.

Le coût de ce service reste toujours inférieur à celui des bases de données propriétaires. La souscription annuelle à EnterpriseDB Advanced Server 8.3 coûte par exemple 5995 \$ par CPU et celle à MySQL Enterprise de 479 € à 3 999 € par serveur et par an. On est loin des 40 000 \$ par CPU demandés par Oracle pour Database 11g Enterprise Edition ou des 29 000 € demandés par Microsoft pour SQL Serveur 2005 Enterprise Edition.

2008 ANNÉE CHARGÉE, SOUS LE SIGNE DU DÉCISIONNEL

2008 sera une année importante et symbolique pour les bases de données Open Source. D'une part, MySQL va s'intro-

duire en bourse, prouvant ainsi la viabilité de ce modèle économique. D'autre part, de nombreux projets décisionnels, de taille et de complexité raisonnable, devraient adopter un SGBD/R libre cette année.

PostgreSQL et Ingres sont particulièrement bien adaptés pour cette problématique, mais MySQL devrait combler son retard avec sa version 5.1 attendue en ce début 2008. Cette nouvelle version introduit en effet certaines fonctionnalités comme « *le support des partitions qui permettra de passer un palier en termes de volume de données dans les projets décisionnels. Et la réplication "row" [NDLR : au niveau des enregistrements] qui permettra de répliquer un cluster et donc de se rapprocher de la technologie Oracle standby* », explique Yann Neuhaus. ■

