

Avant propos

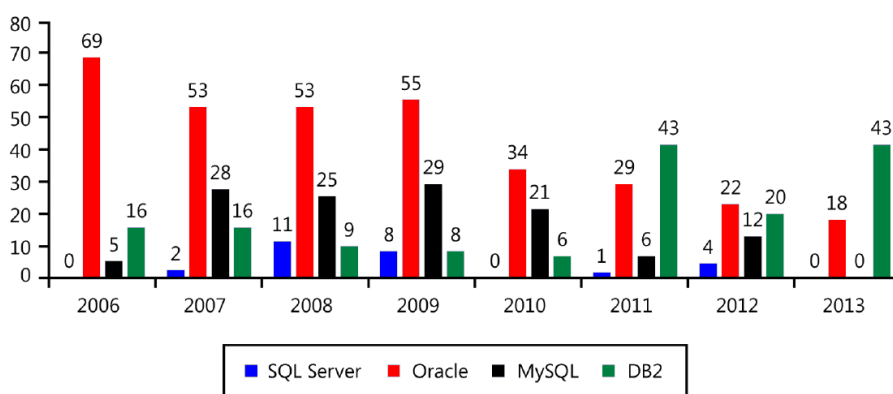
La position de SQL Server dans le monde des SGBDR

SQL Server est l'un des SGBD Relationnel les plus pointus qui soit. Il dépasse aujourd'hui largement son principal concurrent qu'est Oracle sur de nombreux points : **sécurité, fiabilité, performance et couts...**

Peu après la sortie de la version 2005 les experts en **sécurité**¹ avaient pointés du doigt que les importants efforts de refonte du SGBDR avaient apporté un très haut niveau de sécurité devançant largement Oracle. Il est simple de s'en rendre compte en effectuant une requête dans Google concernant les failles de sécurité des deux SGBDR. En tapant les mots clefs adéquats comme NIST² et CVE³ associé au nom du SGBDR, nous trouvons les métriques suivantes : 56 600 résultats pour Oracle et 22 500 pour SQL Server, soit moins de la moitié... C'est encore plus flagrant lorsque l'on effectue cette requête directement dans le moteur de recherche du NIST pour les 3 dernières années : 1 405 cas pour Oracle, contre 191 pour SQL Server... Bref 7 fois plus de bogues parfois critiques, pour Oracle comparés à SQL Server ! Une tendance encore confirmée aujourd'hui par Qualys⁴ qui pointe un nombre de vulnérabilités double dans Oracle que dans SQL Server.

Certains auteurs signalent en outre que la correction des problèmes est très généralement plus lente chez Oracle que chez SQL Server, au moins deux fois plus lente, et certains bogues n'ont pas été corrigés depuis de nombreuses années chez Oracle, alors qu'il s'agit d'une chose impensable pour Microsoft qui a depuis plus de dix ans maintenant mise au point le Microsoft's Secure Development Lifecycle, une méthodologie destinée à aider les développeurs de la firme de Redmond à construire des systèmes plus résilients et naturellement plus fiables.

Et ceci couplé à la très robuste plateforme de test a payé !



Nombre de vulnérabilités recensées par le NIST par an depuis la sortie de SQL Server 2005

¹ Pour l'année 2006, l'analyse de David Litchfield avait montré que Oracle présentait 34 vulnérabilités ayant fait l'objet d'une correction tandis que SQL Server n'en présentait aucune...

² NIST : National Institute of Standards and Technology

³ CVE : Common Vulnerabilities and Exposures

⁴ <http://www.qualys.com/solutions/technology/database/>

La **fiabilité** de Windows associé à SQL Server n'est plus à prouver, n'en déplaisent aux aficionados de Linux. Là aussi de nombreuses études comparatives entreprises sur les coûts d'exploitation⁵ de SQL Server comparés à ceux d'Oracle sous Linux, montrent à la fois la moindre fiabilité de la solution Oracle sous Linux, comparés à la fiabilité accrues ces dernières années du couple de solution Windows / SQL Server. De même les coûts d'exploitation sont bien moindre...

C'est sans doute pour cette raison que la brigade des sapeurs-pompiers de Paris a opté pour SQL Server afin de gérer l'ensemble des systèmes d'alerte avec le logiciel Adagio. C'est aussi pour cela que la DDE du Gard (aujourd'hui DDTM) a confié à SQL Server la gestion des données des cours d'eau du grand delta du Rhône⁶ afin de prévoir les épisodes de crues dont celle de 2003 avait fait plus d'un milliards de dégâts. Le précédent système sous PostgreSQL étant incapable de supporter à la fois la charge, la fiabilité et la haute disponibilité... Une solution de gestion des épisodes de crues aujourd'hui adoptée par une grande partie des SPC⁷ pour d'autres réseaux hydriques et en voie d'introduction dans d'autres pays européens...

Certes on peut arguer qu'Oracle supporte le poids du passé avec un volume de données stockées généralement supérieur à celui de SQL Server, mais les études⁸ ont justement comparées des choses comparables et concordent à affirmer que le cout total d'exploitation est en faveur de la solution de SQL Server et cela en raison de critères objectifs et précis :

- la simplicité d'administration de SQL Server ;
- le faible nombre de combinaison entre les versions du SGBDR et de sa plateforme d'exploitation.

La **simplicité d'administration** est de longtemps un argument de vente des produits Microsoft. Certes il n'est pas possible de tout faire à coup de clic, mais pour bien des tâches simples, les IHM de SQL Server permettent d'obtenir rapidement les informations et administrer facilement les serveurs là où la concurrence exige de se farcir la ligne de commande avec toute la joyeuseté des erreurs de syntaxe qui l'accompagne. Que de temps gagné pour des tâches simples !

Par exemple il est généralement admis qu'une installation d'instance SQL Server nécessite 1,5 heures contre environ 6 pour Oracle...

De plus avec une plateforme système unique, point n'est besoin de commandes à rallonge spécifique à telle ou telle mouture ou version d'OS ! Une simplicité payante qui raccourcit l'écriture des tâches et en optimise l'exécution.

⁵ Le coût d'exploitation est directement lié à l'indisponibilité des données résultant de la fiabilité du système. Des données inutilisables du fait de l'indisponibilité du système coûtent cher à l'entreprise...

⁶ Logiciel « Aquaréal » développé par Synaspe et SQL Spot.

⁷ SPC : Service de Prévision des Crues dépendant des DDT/DDTM

⁸ La première étude sur le genre est celle de Alinean Inc. portant sur la comparaison de Oracle 10 g et SQL Server 2005 dans une centaine d'entreprises sur des volumes comparable, et montrait que le CTO était de 2 847 \$ par an pour SQL Server contre 10 206 \$ par an pour Oracle (publiée en 2006).

La plateforme unique Windows simplifie encore plus la vie lorsque l'on analyse l'étendue des situations de Linux. Loin d'être une richesse, l'énorme quantité des distributions (plus de cent à ce jour, parmi lesquelles RedHat, Debian, Suse, Mandriva, Ubuntu, CentOS...) et des multiples versions des différents Linux complexifient de manière démesurée l'exploitation et la maintenance des systèmes. En effet, il devient difficile d'industrialiser la maintenance lorsque l'on est confronté à plus de 200 versions différentes d'un prétendu même système, comme c'est le cas chez la plupart des grandes entreprises ayant massivement adopté Linux dans des temps passés et aujourd'hui contraints de faire machine arrière dans le but de réduire les coûts !

Quelques récentes études menées en France dans des entreprises autrefois pro linuxiennes et étatiques ont bien montrés les désastres engendrés par une politique absurde du tout free idéologique sans avoir préalablement entrepris des études prospectives pour en connaître les tenants et aboutissants notamment économiques...

Sur le plan des **performances**, SQL Server avait déjà devancé son principal concurrent en matière d'extraction des données, mais pas en ce qui concernait le pur transactionnel. Aujourd'hui l'écart est comblé, SQL Server dépasse Oracle en matière de vitesse transactionnelle avec Hekaton, c'est à dire la technologie « In Memory », initié par SAP et son SGBDR Hana DB. Il est à noter qu'aucun des concurrents du libre n'offre aujourd'hui une telle solution combinant base relationnelle et technologie « In Memory ».

C'est sans doute pour ses performances, et aussi pour son moindre coût d'exploitation que la plupart des grands sites Web marchands de France ont optés pour SQL Server : fnac.com, CDiscount, Ventes Privées...

En comparaison, le site d'annonce faiblement transactionnel « Le Bon Coin » est obligé d'arrêter le services des données de ses serveurs PostGreSQL pour effectuer les principales tâches de maintenance (sauvegarde, réindexation, nettoyage des versions de lignes...) qui s'avéraient incompatible avec la production, même aux heures creuses ! Heureusement sur ce point, SQL Server sait faire en pleine production sans vraiment gêner les utilisateurs concurrents, tant pour la réindexation, que pour la modification de la structure des objets (tables en particulier). Et avec le bon hardware il est même possible de rajouter CPU ou RAM à chaud...

En matière de **coûts**, il y a longtemps que SQL Server est considéré comme un SGBDR « Low Cost » et gagne de nombreux marchés au détriment d'Oracle. C'est en particulier un choix que font de très nombreuses entreprises sur des nouveaux projets, notamment dans la BI, tant le Gartner Group⁹ a montré à quel point Microsoft était en pointe dans la BI depuis plusieurs années... Il n'est qu'à indiquer quelques-unes des réussites françaises dans ce domaine : Bouygues, Havas Média, Conforama, l'Occitane, Véolia, L'Oréal...

Certains croient naïvement que SQL Server n'est qu'un simple Access amélioré... Ce serait faire injure à son ancêtre, Sybase fortement implanté à l'origine dans le monde bancaire... Que les volumes gérables par SQL Server seraient minimes par rapport à ceux d'Oracle. Qu'on en juge... De nombreux acteurs français ont largement dépassés la dizaine de terra octets. C'est le cas de fnac.com,

⁹ À voir par exemple sur ZDnet : <http://www.zdnet.com/gartner-releases-2013-bi-magic-quadrant-7000011264/>

CDiscout, Vente Privées, Essilor... citons aussi le cas de BMW avec plus de 70 To de données en Allemagne...

De nombreuses banques françaises utilisent SQL Server au quotidien pour de très nombreuses bases : Société Générale (plusieurs milliers) Crédit Agricole (plus de 2 000), BNP Paribas (plus de 500)... mais aussi des sociétés de service comme AutoLib (groupe Bolloré) ou Velib (groupe Decaux) ou encore Bouygues...

Quant en matière de prix de vente, il semblerait bien que la concurrence soit aux abois. Certains clients ont obtenu d'Oracle des réductions allant jusqu'à 70% du tarif officiel, là où MS applique des tarifs très encadrés, dont la marge de manœuvre est bloquée à 40% pour les clients les plus pauvres (souvent des administrations) qui achètent des volumes conséquent de licences. Mais même avec un tel rabais, Oracle reste démesurément cher¹⁰ et fait payer au tarif fort le moindre « package » supplémentaire. Qu'on en juge : 5 000 \$ par processeur pour le diagnostic pack ou le Tuning Pack, 11 500 \$ par processeur pour Data Guard (l'équivalent du Mirroring) la compression ou le partitionnement, 17 500 \$ par processeur pour le spatial, 46 000 \$ par processeur pour le moteur OLAP et l'analytique... Alors que tout cela est compris dans la version Enterprise de SQL Server !

Pas étonnant alors qu'Oracle stagne en parts de marché qui sont calculées sur le chiffre d'affaire alors que SQL Server progresse largement en nombre de licences...

Et le libre ?

Il est donc certain que SQL Server possède de nombreux atouts à la fois devant la concurrence commerciale mais aussi devant les SGBDR pseudo ou réellement libre que sont MySQL ou PostgreSQL...

En fait ni MySQL, ni PostgreSQL ne sont de véritables SGBD Relationnels, tous deux échouant aux tests ensemblistes¹¹ comme à celui invoqué par Chris Date dans un célèbre article intitulé « A cure for Madness » et indiquant quel devait être le comportement logique d'un SGBD Relationnel.

Cet échec est la conséquence d'un développement superficiel incapable de prendre en compte la réelle complexité du monde relationnel. C'est assez flagrant quand on se rend compte que tant MySQL que PostgreSQL ne peuvent pas utiliser plus d'un thread pour une même requête s'interdisant ainsi tout parallélisme d'accès ou de manipulation des données. Quant aux aspects de BI ou « In Memory », ces « free bases » n'en sont évidemment pas dotés.

Même si PostgreSQL offre un bon support du SQL¹², il reste très pauvre en outil d'administration ou de diagnostic et catastrophiquement lent dans certaines manipulations d'administration comme par exemple lors des migrations de bases d'un serveur à l'autre¹³...

¹⁰ Entre 2 et 24 fois plus cher selon les configurations par rapport aux tarifs officiellement publiés.

¹¹ Il n'est par exemple pas possible de mettre à jour une colonne clé dans certaines conditions. Voir <http://blog.developpez.com/sqlpro/p10916>

¹² MySQL offre un support très restreint du SQL basé sur une grande partie de la norme de 1992, soit plus de 22 ans de retard ! Pour pallier certains de ses défauts il offre en compensation des fonctions très exotiques loin du fonctionnement normal d'un SGBD Relationnel...

Et au final, aucun des acteurs du libre n'offre une telle richesse de version pour des besoins plus spécifique :

- un moteur relationnel pour stocker et manipuler (SQL) des données des bases transactionnelles (OLTP) ;
- un moteur transactionnel dans le « cloud » (Azure) pour ceux qui ne veulent pas s'embarrasser de la partie purement « système » d'un SGBDR (haute disponibilité en particulier) ;
- un moteur décisionnel pour stocker et manipuler (MDX/DMX) des données des bases analytiques (OLAP) ;
- un ETL intégré hautement multithreadé et super performant pour l'alimentation des DataWarehouse (1 To en 30 minutes...) ;
- un outil de reporting dynamique, sécurisé avec un cache intégré ;
- un outil visuel de manipulation de tableaux croisés dynamique avec « drilling » (PowerPivot) ;
- des extensions pour la gestion de très forts volumes de données non structurées avec HD Insight afin d'aller dans le « Big Data ».

Et finalement le gratuit s'avère cher... Dès que les besoins applicatifs sortent du bas de gamme, le recours au libre nécessite des palliatifs que la différence de licence finit par absorber. Par exemple, dans le cas de la gestion des crues du grand delta du Rhône, il avait été imposé au SPC de la DDTM du Gard de rester dans le libre afin de minimiser les coûts. Mais avec PostGreSQL, il fallait d'énormes ressources pour les serveurs afin de compenser l'absence des vues indexées¹⁴ (donc un coût de serveur démesuré), du matériel, une infrastructure et des logiciels supplémentaires afin de répliquer de manière fiable et cohérente l'ensemble des données de la base à grande distance (de Nîmes à Avignon) dans des tuyaux ne supportant pas plus de 4 mb de bande passante au mieux. Il a donc fallu se battre contre une administration centrale bornée pour imposer SQL Server au détriment PostGreSQL, l'argument massue étant le total hors budget de la solution PostGreSQL avant même le moindre commencement de début d'exploitation...

Il n'est qu'à penser aux pertes de chiffre d'affaire que « Le Bon Coin » se permet en coupant PostGreSQL la nuit pour maintenance, ce qui lui retire la manne que pourrait lui rapporter les annonces des nombreux francophones vivants en horaires décalés, comme c'est le cas des Antilles, de la Polynésie ou de nouvelle Calédonie, mais encore et surtout des canadiens qui seraient heureux sans doute de pouvoir profiter d'un acteur d'annonce aussi populaire !

Des métriques

¹³ Il suffit de quelques millisecondes pour migrer une base de données SQL Server d'une instance à l'autre même quand il s'agit d'une migration évolutive en version (par exemple de 2008 à 2014...).

¹⁴ Certaines tables comptant plusieurs centaines de millions de lignes.

Parmi les métriques record, nos collègues de PASS¹⁵ ont relevé il y a quelques années les chiffres suivants, déjà largement dépassés :

Base de données monolithique la plus grosse	100 To
Table la plus « longue »	1,5 milliards de milliards de lignes
Nombre de transactions le plus élevé	200 000 transactions par secondes ¹⁶
Application ayant le plus gros volume de données	88 Po
Vitesse de chargement des données dans une base	1 To en 30 minutes

Parlons maintenant de quelques projets lourds.

- BWIN : plus de 100 instances de SQL Server abritant quelque 120 To de données sur plus de 1 400 bases, géré par 5 DBA plus un architecte de bases de données... Et sur le serveur principal, plus de 450 000 requêtes SQL par seconde...
- Itaù (Brésil) : BI alimenté par 180 millions de faits par jour. Croissance annuelle de 30 à 50 To.
- Pann-STARRS (Hawaï) : base relationnelle comportant plus de 100 To de données (en 2009) sur un ensemble de 8 serveur fédérés.

Parmi les acteurs connus pour dépasser allègrement les 100 To, il y a Centipede, MySpace, HotMail... et quelque uns ayant franchi la barre du péta octet.

¹⁵ PASS : Professional Association for SQL Server (organisation non lucratives internationales destinées aux utilisateurs de MS SQL Server)

¹⁶ SIB Liquidity Market