

Virtualisez votre Environnement Base de Données

L'impact produit par l'indisponibilité des bases de données

Les entreprises continuent aujourd'hui à se confronter à un problème séculaire : comment permettre l'accès à leurs données critiques, sans interruptions. La non disponibilité des applications peut avoir un impact significatif sur les affaires - perte de revenu, baisse de productivité des salariés, amendes réglementaires et plus mauvais encore, atteinte à l'image de marque de l'entreprise. La plupart des organisations utilisent des serveurs SMP coûteux, isolés ou groupés en cluster pour fournir une certaine disponibilité. Malheureusement ces technologies de "haute disponibilité" ont été conçues pour gérer les pannes, pas une disponibilité 24x7. Elles ne permettent pas d'exécuter la maintenance sur leurs bases de données sans l'arrêt de l'exploitation. En conséquence, le service informatique est forcé de maintenir des systèmes actifs-passifs chers et le recours à la planification d'arrêt d'exploitation pour garder les bases de données à jour.

La principale motivation de la virtualisation est la réduction des coûts. C'est pourquoi les fournisseurs de logiciels et d'infrastructure se sont attaqués au domaine extrêmement coûteux des serveurs et du stockage. Maintenant, il existe un autre domaine banalisé dans lequel les entreprises dépensent des sommes très importantes : les bases de données.

Mais, force est de constater que très peu de logiciels sont présents sur le marché pour résoudre ce problème complexe : comment réduire les coûts des bases de données qui deviennent de plus en plus grosses et difficiles à maintenir d'année en année.

La virtualisation des bases de données, permet donc de réduire les coûts de maintenance du nombre important des data marts, datawarehouses et toutes les bases de production en permettant leur gestion sur des hardware banalisés de type commodité et en réduisant les processus de copie des données, coûteux en ressources machine.

Les utilisateurs n'accèdent plus aux données brutes, les requêtes exécutées sur une instance locale, peuvent devenir logiques et être exécutées sur des bases de données multiples dispersées sur le réseau. Les DBA n'ont plus besoin de déplacer autant de données entre les bases rendant par exemple le chargement de données avec des ETL moins lourds.

De plus la virtualisation des données fournit une plus grande souplesse dans le déploiement des serveurs de données et une réponse rapide aux changements, par l'extensibilité (scalability) apportée par cette nouvelle technologie de pointe.

Pour ceux qui pensent que la virtualisation des bases de données est une couche supplémentaire qui va créer une surcharge au niveau de l'administration et de la maintenance, l'expérience a montré que les autres domaines 'virtualisés' ont généré des gains substantiels en terme de souplesse et de potentiel de consolidation.

Il n'est pas question de supprimer les unités de stockage et les serveurs de données, en effet les données augmentent de façon exponentielle (selon les analystes, leur volume double tous les 18 mois). Ce qui se produit avec la virtualisation, c'est la transparence du stockage et des serveurs de bases, qui deviennent logiques, augmentant le potentiel de consolidation des données, leur accès unifié, et le potentiel de remplacement des équipements âgés par de l'infrastructure de type commodité donc moins chère.

Une des préoccupations des DBA concerne la performance des bases de données virtualisées, ainsi le problème des E/S et des locks est posé. En réalité, c'est le contraire que l'on observe car les données sont stockées sur des serveurs de bases de données multiples permettant des mécanismes de gestion de charge 'load balancing'. Les requêtes sont envoyées au serveur ayant la meilleure disponibilité ce qui augmente les performances globales.

On a vu que la virtualisation apporte un accès transparent aux données quelque soit l'endroit physique de stockage. Elles existent donc dans plusieurs bases à la fois, leur mise à jour est prise en charge de manière synchrone ou asynchrone en fonction de leur criticité et de la performance attendue. En mode synchrone l'application attend que le dernier serveur ait fait sa mise à jour. Ceci peut avoir un impacte sur les performances si l'application fait en majorité des écritures, en revanche le contraire se produit si l'application fait majoritairement des lectures, ce qui est souvent le cas. Le load balancing est le facteur qui équilibre les performances.

Il faut aussi comprendre que la virtualisation des bases de données ne découle pas de la virtualisation du support. Par exemple, ce n'est pas parce que la base est stockée sur une machine virtuelle ou sur une plateforme virtuelle qu'elle est elle-même automatiquement virtualisée. C'est un Service qui fédère les données et isole l'application de l'endroit où elles sont stockées, le 'load balancing' permettant d'exploiter au mieux les serveurs de données.

La gestion contrôlée de la redondance des données apporte un avantage supplémentaire : la haute disponibilité des données, à condition que les serveurs de bases de données soient de type Actif-Actif, si un serveur de bases tombe en panne les données sont toujours disponibles. De même dans le cas d'une maintenance d'un serveur ou d'une base de données, les applications ne sont pas impactées pour autant.

En conclusion, la virtualisation de bases de données est le nouveau domaine où des économies importantes peuvent être réalisées en apportant une souplesse de gestion plus importante ainsi qu'une disponibilité quasi-totale des données sans pour autant pénaliser les performances des applications.



Plus d'informations

Sur le site d'EGEOSOFT, distributeur de la société Xkoto
www.egeosoft.com
+33 1 44 54 67 16
ou www.xkoto.com



GRIDSCALE and xkoto are trademarks or registered trademarks of xkoto Inc. All other names may be the trademarks or registered trademarks of their respective owners. All rights reserved. ©2005-2008 xkoto Inc. Item: GSSQL-2-EN-20080606