

Solutions de Business Intelligence avec Alcade et les Technologies InterSystems

Un livre blanc ALCADÉ

janvier 2006

Introduction

La tendance actuelle, concernant les solutions de business intelligence, est de réduire les temps d'accès aux données pour traiter et analyser des informations toujours plus à jour. L'aboutissement de cette tendance est la création d'applications d'un nouveau type. Les nouvelles applications de *business intelligence* travaillent sur des données transactionnelles et les analysent en temps réel. Les solutions de *Business Activity Monitoring* font de même pour les systèmes intégrés.

La base de données objet InterSystems Caché® a une architecture de stockage des données multidimensionnelle qui apporte des performances jusqu'à 5 fois plus élevées qu'avec les bases de données relationnelles. Ces performances élevées font d' InterSystems Caché® LA base de données idéale pour les solutions de *Datawarehouse* traditionnelle, comme pour des analyses en temps réel sur des données transactionnelles en utilisant les index *bitmap*.

InterSystems Caché® est la seule base de données capable de gérer les index *bitmap* en transaction

La plate-forme d'intégration unifiée universelle InterSystems Ensemble® permet la même analyse en temps réel, sur des systèmes intégrés, permettant de bâtir facilement des solutions de *Business Activity Monitoring*.

La *Business Intelligence* aujourd'hui

Les données traitées chaque jour par l'ensemble des systèmes d'information d'une entreprise revêtent une importance tactique et stratégique considérable, dans la mesure où elles peuvent être traitées, analysées, et transformées rapidement en information utile. C'est ce que font les applications de *business intelligence*. Nombreuses sont les entreprises qui mettent en œuvre d'importants moyens humains et financiers pour bâtir, maintenir, et améliorer ces applications.

L'architecture la plus répandue pour les applications de *business intelligence* est fondée sur une base de données relationnelle utilisée comme *datawarehouse*. Cette base est régulièrement remise à jour en mode batch à partir des données issues d'une ou plusieurs bases transactionnelles. Les données ainsi stockées peuvent être traitées et analysées par des outils de reporting, utilisant le datamining et les requêtes appropriées.

Mais les utilisateurs ont rapidement découvert les limites des *datawarhouses*. L'une est leur faible niveau de performance lorsqu'il s'agit de manipuler des données complexes. Une autre est qu'en fonction de la fréquence des mises à jour, les données peuvent être vieilles de plusieurs heures, voire de plusieurs jours.

En conséquence, les applications de *business intelligence* tendent à reposer sur des *datawarehouses* plus rapides et plus fréquemment mis à jour. La suite logique de cette tendance est de créer des applications capables des mêmes analyses, mais directement sur les systèmes en transaction, en temps réel.

L'analyse en temps réel des données est également la base des systèmes intégrés. Les solutions de *Business Activity Monitoring* (BAM) proposent généralement des tableaux de bord regroupant en temps réel les différents indicateurs d'activité stratégiques pour l'entreprise.

Ce livre blanc montre comment la base de données objet InterSystems Caché® peut être utilisée pour bâtir des applications de business intelligence et des *datawarehouses* à très hautes performances, des systèmes d'information capables de délivrer des informations en temps réel, sur des données transactionnelles. Ce livre présente également les fonctionnalités qui font de la plate-forme d'intégration unifiée InterSystems Ensemble® l'outil idéal pour le *Business Activity Monitoring* et les systèmes intégrés.

InterSystems Caché® utilisé comme *Datawarehouse* à très hautes performances

Bien qu'offrant toutes les fonctionnalités d'une base de données relationnelle, InterSystems Caché® est bien plus que cela. Les données y sont stockées sous forme de tableaux multidimensionnels. Ce mode de stockage est bien plus performant pour accéder aux données (qui plus est pour des données complexes), que la structure tabulaire à deux dimensions des bases de données relationnelles. Parce que la base objet InterSystems Caché® n'est pas soumise aux règles de « jointures » et de « relations entre les tables » qui caractérisent les bases de données relationnelles, elle consomme moins de *process* et répond beaucoup plus rapidement aux requêtes.

Bien que InterSystems Caché® ne soit pas une base de données relationnelle, elle accepte les accès ODBC (et JDBC) utilisant SQL comme langage de requêtes. Cela tient au fait que, bien que stockant les données sous forme de tableaux multidimensionnels, InterSystems Caché® offre une projection SQL qui présente les données comme si elles étaient stockées sous forme de tables relationnelles. De fait, InterSystems Caché® est ouvert à l'ensemble des outils d'analyse et de reporting initialement conçus pour les bases relationnelles. Les tests comparatifs réalisés sur des bases en production sur sites clients montrent que les temps de réponse aux requêtes SQL d'InterSystems Caché® sont en moyenne 5 fois plus faibles qu'avec toute autre base de données relationnelle du marché. InterSystems Caché® est également 5 fois plus rapide lors de la

mise à jour des données d'un *datawarehouse* par procédures batch.

Des performances bien supérieures encore sont facilement obtenues en utilisant la puissance de l'accès direct aux données multidimensionnelles offert par InterSystems Caché®. Un certain nombre d'outils d'analyse et de reporting du marché utilisent déjà en natif l'accès direct InterSystems Caché®.

Non seulement un *datawarehouse* basé sur InterSystems Caché® est plus rapide, mais encore, il utilise moins de ressources matérielles qu'une base de données relationnelle servant les mêmes données. Grâce à l'organisation multidimensionnelle de ses données, InterSystems Caché® utilise la moitié, voire le tiers de l'espace disque utilisé par un SGBD relationnel. De plus, le fonctionnement d'InterSystems Caché® ne requiert pas de « réglage » particulier, et son administration est réduite au minimum.

Traiter des données plus fraîches

Certaines applications de business intelligence tolèrent une relative obsolescence des données. Les analyses au mois le mois peuvent supporter des mises à jour espacées, et des résultats de requêtes relativement longs à obtenir. Mais d'autres applications, doivent pouvoir traiter des données plus fraîches, sur un mode proche du temps réel.

Un des moyens les plus souvent utilisés pour limiter l'obsolescence des données est d'utiliser un serveur miroir (*shadow server*). Lorsque les données sont mises à jour sur le serveur primaire (le serveur opérationnel), les transactions sont enregistrées dans un fichier journal. Un ou plusieurs serveurs miroirs répercutent les changements dans une base de données dupliquée. Les requêtes sont adressées aux serveurs de bases dupliquées, évitant ainsi la dégradation des performances du serveur de production. L'obsolescence des données est limitée au temps nécessaire pour lire et appliquer les transactions journalisées. En fonction du nombre d'accès

auquel le serveur miroir est soumis, l'obsolescence des données peut être de plusieurs secondes.

InterSystems Caché® supporte la technologie de *shadowing*. Ses très hautes performances permettent de minimiser le temps nécessaire à la mise à jour des données, comme celui nécessaire à la réponse aux requêtes.

InterSystems Caché® et l'analyse « temps réel »

Dans certains cas, même un délai de latence de quelques secondes peut être trop long. Il est parfois indispensable que l'analyse porte sur des données en temps réel. En ce cas, il est vital que le processus de requêtes soit extrêmement rapide, et ne dégrade pas les performances opérationnelles de l'ensemble du système.

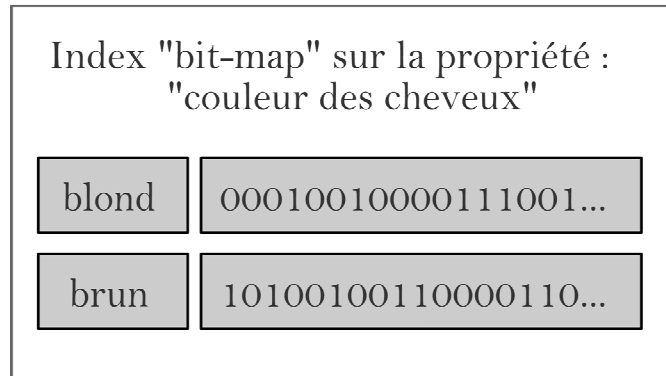
La solution pour tout système de base de données pour améliorer les temps de réponses est de créer des index sur les fichiers. InterSystems Caché® offre aux développeurs la possibilité de créer aussi bien des index traditionnels que des index *bitmap*.

Dans les systèmes relationnels, les index sont des listes d'ID correspondant à une propriété (colonne de table dans les SGBDR).

Index traditionnels sur la propriété : "couleur des cheveux"	
blond	4,7,18,25,32,40,95,103,...
brun	1,8,16,28,29,36,47,75,...

Les index *bitmap* sont souvent utilisés dans les solutions de *datawarehouse* car ils améliorent de façon très significative les temps de réponse, notamment lorsque l'on traite de grands volumes de données. Dans un index *bitmap*,

une valeur est enregistrée sous forme de chaîne de bit représentant chaque enregistrement. Le bit prend la valeur « 1 » si l'enregistrement possède la propriété, « 0 » si il ne la possède pas.



L'avantage des index *bitmap* est que les requêtes s'exécutent en recourant uniquement à des opérateurs booléens (AND, OR) sur les index, permettant de déterminer avec rapidité quels sont les enregistrement correspondant à la recherche, sans parcourir l'ensemble de la base de données. Dans le cas de requêtes complexes, l'utilisation d'index *bitmap* peut diminuer les temps de réponses d'un facteur 100, voire plus.

Cependant, les index *bitmap* sont difficiles à maintenir lorsque les enregistrements sont modifiés ou supprimés fréquemment (bases en transaction). C'est pourquoi ils sont utilisés uniquement sur des bases de données en lecture seule comme le sont les *datawarehouses*. Enfin, les index *bitmap* peuvent occuper un espace disque important puisque chaque valeur possible de l'index contient un bit (0 ou 1) pour chaque enregistrement de la base.

InterSystems Caché® est le seul SGBD capable de gérer les index *bitmap* sur des bases en transaction (lecture / écriture / suppression).

La structure multidimensionnelle, et les techniques de compression de données font que les index *bitmap* d'InterSystems Caché® occupent moins d'espace disque, et sont plus rapide à maintenir qu'avec tout autre système de base de données. Avec InterSystems Caché®, la maintenance des index *bitmap* ne prend pas plus de temps ni de ressources que si il

s'agissait d'index traditionnels. Ils sont donc utilisés de façon massive sur des bases de données transactionnelles.

L'utilisation d'index *bitmap* sur des bases de données en transaction rend possible la création d'applications de *business intelligence* capables d'analyser des données opérationnelles en environnement "temps réel".

InterSystems Ensemble® : *Business Activity Monitoring* et l'intégration d'applications

La capacité à analyser les données en temps réel est également une contrainte des systèmes intégrés proposant des fonctions de *Business Activity Monitoring* (BAM). Dans le cas d'application BAM, les informations sont collectées en temps réel en provenance de sources de données hétérogènes.

La plate-forme d'intégration universelle InterSystems Ensemble® embarque une solution de stockage des données enregistrant tous les messages passant par le système intégré. Cet "entrepôt de messages" est fondé sur une structure multidimensionnelle des données semblable à celle d'InterSystems Caché®, et offrant les mêmes performances en terme de temps de réponse. InterSystems Ensemble® analyse les données de son "entrepôt de messages" en temps réel, ce qui permet de réaliser des application BAM d'une puissance inégalée.

L'environnement d'intégration InterSystems Ensemble® inclus des fonctionnalités permettant de développer plus rapidement des applications BAM. La technologie objet embarquée dans InterSystems Ensemble®, et les assistants du studio de développement rendent simple la création d'indicateurs graphiques qui montrent, ou calculent l'information présentée en garantissant un rafraîchissement automatique rapide des données issues de l'ensemble des systèmes intégrés. Ces informations sont automatiquement associées à des graphes, des cadrans, ou tout indicateur graphique permettant de diffuser des tableaux de bord de management fondés sur une technologie web. Les indicateurs sont couplés si besoin est à des procédures de retour

d'informations, agissant directement sur les applications intégrées. Exemple : Il est possible d'adresser une instruction à une application en cas de dépassement d'une valeur limite constatée par un indicateur graphique.

Conclusion

Les produits InterSystems Caché® (base de données multidimensionnelle à très hautes performances) et InterSystems Ensemble® (plate-forme d'intégration universelle) utilisent tous deux une architecture de données multidimensionnelle qui convient particulièrement aux solutions de *business intelligence*. Les applications de *datawarehouse* traditionnelles bénéficient des temps de réponses incroyablement rapides d'InterSystems Caché®. De telles performances permettent une analyse en temps réel sur des données hautement transactionnelles, ouvrant de nouveaux champs d'applications aux solutions de *business intelligence*. Avec les technologies InterSystems®, les applications de *business intelligence* sont indépendantes et capables d'analyser et d'interagir en temps réel, permettant aux entreprises de tirer pleinement profit des données traitées quotidiennement.

InterSystems Ensemble® permet une analyse en temps réel des messages et autres données transitant à travers un système intégré. Avec InterSystems Ensemble®, il devient facile de créer des outils de contrôle et des mécanismes de correction automatisés utilisés par les applications de *Business Activity Monitoring*.

ALCADE est expert en solutions fondées sur les technologies InterSystems®. Que ce soit pour répondre à des besoins de performances de bases de données, ou des besoins de développement rapide d'Application d'Intégration d'Entreprise, ALCADÉ est capable de proposer des solutions adaptées en terme de délai de réalisation et de budget.

A propos d'ALCADE

ALCADE est intégrateur d'applications d'entreprise, expert en gestion et administration de bases de données.

ALCADE réalise des applications de gestion en environnement web (Internet – Intranet – Extranet) à très hautes performances.

ALCADE dispose d'une expertise reconnue au niveau international sur les technologies InterSystems Caché® et InterSystems Ensemble®.

ALCADE

BP 244 - 310, rue du Vallon
06905 SOPHIA ANTIPOLIS
CEDEX

Tél : +33 (0)4 93 74 66 06
Fax : +33 (0)4 93 74 04 00

Email : infos@alcade.com

www.alcade.com