

Informix.
software

Informix Newsletter

2nd trimestre 2011

User Group Informix France

Bienvenue dans la Newsletter du User Group Informix France

Au programme

Editorial	3
Retour sur l'IIUG Informix Conference 2011	5
WebTip: Espace français sur iiug.org et informix.fr réactivé	6
Technical Tip: IMPLICIT_PDQ.....	6
Technical Tip: L'authentification interne des utilisateurs.....	8
Technical Tip: Nouvelles fonctionnalités en 11.70.xC2	10
Technical Tip: Nouvelles fonctionnalités en 11.70.xC3	13
Etude de cas : Objectif performance.....	16
Derniers articles	21
Vidéos	21
Liens Utiles.....	21
Informix blogs	21
Abonnement / Annulation / Avis.....	22
Les contributeurs de ce numéro	22

Editorial



Nous sommes heureux de vous envoyer le second numéro de la newsletter du *User Group Informix France* pour 2011.

Voilà 10 ans que IBM a fait l'acquisition d'un acteur majeur dans le marché concurrentiel des bases de données d'entreprises.

IBM continue à faire évoluer la gamme des produits Informix avec :

- IBM Informix Dynamix Server V11.70
- IBM Informix Ultimate Warehouse Edition
- IBM Informix Genero

IBM Informix est un produit stratégique dans l'offre Information Management.

IBM propose 5 éditions différentes pour répondre à vos exigences allant de la seule édition gratuite sur le marché vous permettant de gérer des bases de données illimitées en taille jusqu'à l'édition Ultimate qui représente le nec plus ultra en terme de fonctionnalités pour répondre à toutes vos demandes et besoins adaptés aux technologies les plus avancées d'aujourd'hui et de demain.

Quatre nouveaux axes d'innovations pour cette année:

- Informix Warehouse : 100 fois plus rapide que le moteur Informix IDS actuel
- Flexible Grid : cluster hétérogène d'entreprise (des centaines voire milliers de serveurs)
- Genero : modernisez vos applications 4GL et allez vers le Web et les interfaces mobiles
- TimeSeries : la solution la plus performante du marché

IBM Informix reste à jamais un leader à la pointe de la technologie dans le monde des bases de données haut de gamme à la fois au niveau des fonctionnalités, de la performance et en termes d'ouverture aux outils clients du marché et en particulier à l'Open Source.

La conférence annuelle 2011 de l'IIUG était un succès avec plus de 130 sessions et ateliers pratiques.

Continuez à investir et faire évoluer vos développements autour d'Informix.

Nous tenons à vous remercier d'être venu nombreux au dernier évènement de l'UGIF le 4 avril 2011 et espérons vous voir encore plus nombreux à notre prochain évènement à l'automne prochain.

Bonne Lecture !

Khaled Bentebal - *Président de l'UGIF*

Olivier Bourdin - *Editeur de la Newsletter de l'UGIF*

Retour sur l'IIUG Informix Conference 2011



Du 15 au 18 mai s'est tenue à Overland Park, Kansas la conférence annuelle des utilisateurs Informix.

Ce fut une grande conférence très riche en information et en nouveautés. Une conférence annuelle pour couronner une année pleine d'annonces: la sortie de la version majeure 11.70.xC1 de Informix Dynamic Server au mois d'octobre 2010, un nouveau produit innovateur dans le domaine décisionnel, sans aucune concurrence sur le marché, Informix Ultimate Warehouse Edition le 29 mars 2011 et IBM Informix Genero.

Que de sessions et d'informations :

- 3 Sessions plénières
- 2 Tables rondes : technologique et stratégique
- 24 Sessions de type tutoriel très pratique
- 95 Sessions techniques
- 19 Ateliers pratiques (Hands On Labs)
- 7 Certifications gratuites sur Informix V11.50 et 11.70 pendant 3 jours

Le point culminant de la conférence était une session plénière animée par Dr David Ferrucci. Dr Ferrucci est le chercheur et l'investigateur principal du projet Watson / Jeopardy chez IBM. Il est le patron d'une équipe de 28 chercheurs et développeurs dans le domaine du traitement des langages naturels et principalement du projet Watson depuis 2007.

Les 2 autres sessions plénières étaient animées par :

- Jerry Keesee, *Director of Informix Database Development – IBM*
- Dr. Arvind Krishna, *General Manger, Information Management – IBM*
- Rob Thomas, *Vice President, Business Development, Information Management Software Division – IBM*

Sans oublier un hall d'exposition de plusieurs sponsors présentant leurs solutions autour de la base de données Informix.

Les présentations sont disponibles en téléchargement sur le site www.iiug.org dans la zone réservée aux membres

WebTip: Espace français sur iiug.org et informix.fr réactivé

Désormais vous pouvez retrouver sur le site de l'IIUG les Newsletters du groupe des utilisateurs Informix France à l'adresse : <http://www.iiug.org/intl/fra/> . Cette page vous permet en outre d'accéder au forum de discussions IIUG francophone.

Le domaine Informix France (<http://informix.fr>) fonctionne à nouveau grâce à Jean-Georges Perrin. Retrouvez sur ce site les dernières nouvelles sur Informix.



Accueil À la Une Actualités News Jobs Contact

IBM Informix Warehouse Accelerator – La performance c'est tout !

9 juin 2011 | Auteur [Journal Informix](#)

Imaginez que vos analystes de marché, vos responsables d'entreprise recevant leurs rapports en quelques minutes au lieu de quelques heures. Imaginez l'amélioration des capacités d'analyse et de prévision. Imaginez les requêtes de vos « datawarehouse » s'exécutant plus rapidement sans avoir constamment à les surveiller et les optimiser. Imaginez faire tout cela sans avoir à construire de cubes, de tables de résumés, d'index, de statistiques ou d'implémenter des stratégies de partitionnement. Après des tests exhaustifs de « Informix Warehouse Accelerator », Ashutosh Khunte, architecte sénior de base de données chez Sketchers affirme : " Avant l'utilisation de Informix Warehouse Accelerator, les requêtes complexes d'inventaire ou d'analyse de vente sur les magasins de l'entreprise contenant plus d'un milliard de rangées prenaient jusqu'à 45 minutes pour s'exécuter. Lorsque ces mêmes requêtes ont été passées sur Informix Warehouse Accelerator, elles se sont terminées en 2 à 4... ([lire la suite : version PDF](#))

- [Télécharger la plaquette PDF « IBM Informix Warehouse Accelerator »](#)

Publié dans À la Une | Tags : IBM, Informix, Warehouse | [Aucun commentaire »](#)

Recherche

Recherche

IIUG

<http://www.iiug.org>

À la Une

- [IBM Informix Warehouse Accelerator – La performance c'est tout !](#)
- [Proof of Technology Informix IDS 11.70 – Jeudi 16 juin 2011](#)
- [Informix Genero & Ultimate Warehouse Edition](#)

Actualités

- [Des nouvelles du front](#)
- [IDS enfin sur facebook](#)

Technical Tip: IMPLICIT_PDQ

Le PDQ (Parallel Database Query) est un mécanisme qui permet d'obtenir le maximum de performance dans les environnements décisionnel en divisant les requêtes en multiples tâches de scan, jointure, tri, groupe et agrégation qui s'exécutent en parallèle. L'échange de données se fait de manière interne entre les threads producteurs et les threads consommateurs.

Le paramètre *PDQPRIORITY* du fichier *ONCONFIG*, ou sous forme de variable d'environnement, ou de commande SQL détermine le degré de parallélisme pour une requête.

Lorsque la valeur de *PDQPRIORITY* est OFF (0), le PDQ est désactivé et les requêtes ne sont pas parallélisées.

Lorsque la valeur de *PDQPRIORITY* est supérieur à 0, Le MGM (Memory Grant Manager) contrôle et réserve les ressources utilisées par les requêtes dont l'exécution est parallélisée, comme le nombre simultané de requêtes parallèles, le nombre de scan threads pour chaque requête parallèle, le nombre de PDQ threads, ainsi que la quantité de mémoire et CPU disponible pour chaque requête.

Depuis IDS 11.70.xC1 l'activation du PDQ peut se faire de manière implicite.

Il est désormais possible que IDS utilise la priorité PDQ, même en l'absence explicite du paramétrage de *PDQPRIORITY*.

Il est alors possible d'activer le PDQ implicite à l'aide d'une nouvelle variable de session :

- *set environment IMPLICIT_PDQ on/off/[0-100]*

Les plans d'exécution de requêtes STAR JOIN nécessite que la priorité PDQ soit positionnée.

Lorsque la variable de session *IMPLICIT_PDQ* est positionné à 'on' le serveur IDS utilise PDQ indépendamment de la valeur de *PDQPRIORITY* :

- L'optimiseur considère les plans d'exécution de type STAR JOIN même lorsque *PDQPRIORITY* n'est pas explicitement positionné.
- Si l'optimiseur choisi un plan STAR JOIN, le PDQ est implicitement positionné pour permettre l'exécution d'un STAR JOIN.
- Pour les plans qui ne sont pas de type STAR JOIN, PDQ est implicitement positionné pour les requêtes qui respectent certains seuils, afin d'éviter d'utiliser PDQ pour les petites requêtes :
 - Si le coût total d'une requête est supérieur ou égal à 5000, le PDQ est positionné à la fois pour les scans parallèles et jointures parallèles.
 - Si le coût total d'une requête est compris entre 1000 et 5000, le PDQ est positionné seulement pour les scans parallèles.
 - Si le coût total d'une requête est inférieur à 5000, le PDQ n'est pas implicitement positionné.

Par ailleurs, lorsque PDQ est implicitement positionné, le serveur IDS calcule la quantité de mémoire pour chaque thread consommateur, et estime le maximum de mémoire nécessaire à une requête (en se basant sur les estimations de l'optimiseur), afin de positionner le PDQ de sorte à exécuter la requête de manière efficace. Le positionnement explicite du PDQ est ignoré dans ce cas.

Une 2^{ème} variable de session est supportée :

- *set environment BOUND_IMPL_PDQ on/off*

Lorsque la variable de session *BOUND_IMPL_PDQ* est positionnée, le serveur IDS utilise le paramétrage de la priorité PDQ explicite, comme une limite supérieure pour calculer la priorité PDQ implicite. Dans ce cas, si le PDQ implicite calculé est plus petit que le PDQ explicite de la session, le PDQ implicite sera utilisé pour la requête. Cependant, si le PDQ implicite calculé est plus grand que le PDQ explicite de la session, le PDQ explicite de la session sera utilisé pour la requête. Cela permet de fournir une manière de positionner une limite supérieure sur la priorité PDQ implicite calculé par le serveur.

Ces nouvelles commandes SQL « *set environment* » peuvent être utilisées dans les procédures *sysdbopen()* afin que cette nouvelle fonctionnalité soit incorporée sans changement dans le code des applicatifs.

Technical Tip: L'authentification interne des utilisateurs

Avec la 11.70.xC2, la gestion des utilisateurs subit des changements importants. Grâce à l'authentification interne des utilisateurs, il n'est plus nécessaire de dépendre du système d'exploitation ni d'un système externe tel qu'un serveur LDAP pour l'authentification. Les raisons sont multiples pour ne pas vouloir enregistrer de nombreux utilisateurs dans le système d'exploitation qui n'ont d'autres objectifs que de se connecter à la base de données avec, entre autres, une séparation claire des fonctions de l'administrateur système et du DBA. Par ailleurs, toutes les sociétés ne veulent pas forcément installer un système à base de LDAP ou Kerberos, souvent complexe à mettre en œuvre.

La notion des « internal users » s'appuie sur le mécanisme des « mapped users » introduit dans la version précédente: 11.70.xC1, mais en supprimant la nécessité de passer par un système d'authentification externe via PAM (Plugable Authentication Module) ou SSO (Single Sign-On). Le DBA crée des utilisateurs propre à IDS grâce à la commande CREATE USER. Les enregistrements correspondants sont insérés dans la base de données sysuser. Ces utilisateurs n'existant pas au niveau du système d'exploitation, il faut les « mapper » sur un (ou plusieurs) utilisateur existant au niveau du système d'exploitation, appelé « surrogate users » ; l'utilisateur interne sera donc traité en tant que l'utilisateur système. Il est aussi possible de définir des groupes qui auront les mêmes propriétés que des groupes du système d'exploitation.

La mise en œuvre de cette fonctionnalité d'authentification est assurée par un certain nombre de nouveaux ordres SQL :

- CREATE USER
- DROP USER
- ALTER USER
- RENAME USER
- SET USER PASSWORD

Prenons 2 machines (cible et distante) avec chacune une instance de définie et une base de démonstration stores_demo créée sur la machine cible. La machine cible et la machine distante ont chacune les informations nécessaires dans le sqlhosts pour communiquer.

Dans un premier temps, sur la machine cible, nous allons créer un utilisateur interne ifx_int1 qui sera mappé sur l'utilisateur créé au niveau du système d'exploitation : ifx_guest.

```
CREATE USER ifx_int1 WITH PASSWORD "abcdefgh" PROPERTIES USER 'ifx_guest';
```

Il ne faut bien sûr pas oublier de positionner le USERMAPPING à une valeur autre que OFF ; dans le cadre de cette démonstration, nous nous contenterons de la valeur BASIC qui permet de faire fonctionner le USERMAPPING sans pour autant donner des droits administrateurs aux utilisateurs même s'ils sont « mappés » sur des utilisateurs ayant des droits administrateur.

A partir de machine distante, sous un utilisateur quelconque (par exemple « xyz »), nous allons pouvoir passer dans le menu « Connection » de dbaccess , et choisir l'informixserver correspondant à l'instance ou nous avons créé ifx_int1. Au prompt CONNECT, nous tapons ifx_int1 et au prompt PASSWORD

nous tapons abcdefgh. dbaccess nous invite maintenant à choisir la base de donnée (stores_demo) sur laquelle nous souhaitons travailler; la connexion à la machine cible a réussi.

A partir de là, d'autres utilisateurs peuvent être ajoutés. Et à l'aide des autres ordres SQL, cités précédemment, le DBA est en mesure de gérer ses utilisateurs.

Quelques outils qui peuvent se révéler utiles :

```
CREATE DEFAULT USER WITH PROPERTIES USER 'guest' ;
```

Cette commande permet de créer un « surrogate user » par défaut qui sera utilisé pour tous les utilisateurs internes même s'ils n'y ont pas été spécifiquement associés.

Comme dans le cas d'un système d'exploitation, le DBA peut souhaiter désactiver l'accès à son instance pour un utilisateur sans toutefois vouloir le supprimer ; il suffit de le verrouiller.

```
ALTER USER ifx_int1 ACCOUNT LOCK ;
```

Et la commande inverse:

```
ALTER USER ifx_int1 ACCOUNT UNLOCK;
```

Technical Tip: Nouvelles fonctionnalités en 11.70.xC2

Cet article a pour but de vous apporter une vue des nouvelles fonctionnalités introduites dans la version IBM Informix 11.70.xC1.

Private Instance.

La procédure d'installation d'Informix a été améliorée de sorte que vous pouvez installer les produits sans les privilèges root. Ce changement rend plus facile l'installation d'Informix dans des environnements où les privilèges root sont limités ou lorsque le produit fait partie d'une solution embarquée.

Lorsque vous installez le produit sans les privilèges root, le compte utilisateur qui effectue l'installation devient l'administrateur du serveur de bases de données (DBSA) pour cette installation. Vous pouvez alors copier l'installation non-root et la déployer sur d'autres ordinateurs dans le cadre de serveurs de base embarqués.

Une installation non-root ne supporte pas certaines fonctionnalités majeures telles que Entreprise Replication, les connexions distribuées, et la haute disponibilité des clusters.

Déploiement d'Informix via une image RPM sur Linux

Sur les systèmes d'exploitation Linux où vous avez installé rpmbuild, vous pouvez créer une image RPM pour simplifier la redistribution sur plusieurs ordinateurs.

Lorsque vous installez Informix, vous pouvez générer une image RPM Package Manager en excluant des modules du serveur de bases de données et des produits client que vous ne prévoyez pas d'utiliser. Vous pouvez ensuite déployer cette image RPM sur plusieurs ordinateurs Linux qui supportent RPM.

En éliminant sélectivement les fonctionnalités que vous n'avez pas besoin, vous pouvez réduire la taille de l'image distribuable.

New Event Alarms

La fonctionnalité d'Event Alarm introduite en 11.70.xC1 se voit ajouter deux nouvelles class ID.

L'Event Alarm 81 indique la corruption journal logique.

L'Event Alarm 83 indique qu'un serveur SDS ne peut devenir un serveur primaire, car le serveur primaire est toujours actif.

Nouvelles commandes dans le SQL API

L'API SQL reçoit de nouvelles commandes afin d'augmenter sa capacité d'administrer le serveur de bases de données. Les arguments suivants sont désormais disponibles avec fonctions de l'admin SQL API, task() et admin ():

- **create database.** Cet argument crée une base de données et est équivalent à l'instruction CREATE DATABASE.
- **drop database.** Cet argument supprime une base de données et est équivalent à l'instruction DROP DATABASE. Cette fonction supprime la base de données entière, y compris tous les catalogues système, les objets et les données.
- **ontape archive.** Cet argument invoque l'utilitaire ontape pour lancer une sauvegarde.
- **onbar.** Cet argument permet d'exécuter des options spécifiques de l'utilitaire onbar.
- **onsmsync.** Cet argument invoque l'utilitaire onsmsync afin de synchroniser la base de données sysutils, le storage manager, et l'emergency boot file ixbar.

Mode « case-insensitive » pour les chaînes de caractères NCHAR et NVARCHAR

Dans les versions précédentes d'IBM Informix, les chaînes stockées dans toutes les bases de données Informix ont été traitées comme « case-sensible » par les opérations de base de données. Par exemple, une requête avec une clause where recherchant la chaîne "McDavid" retournera uniquement "McDavid" et non pas "mcdavid", "Mcdavid", ou "MCDAVID". Les opérations visant à ignorer le « case » des chaînes de caractères exigent un index BTS ou un index fonctionnel pour chacune des requêtes.

Dans cette version une base de données est toujours créée, comme « case-insensitive » par défaut. Cependant, vous pouvez utiliser l'option NLSCASE INSENSITIVE dans l'instruction CREATE DATABASE pour créer une base de données qui ignorera le « case » des chaînes de caractères. Par exemple, l'interrogation avec la chaîne "McDavid" retournera les rows contenant "mcdavid", "McDavid", "MCDAVID", et "Mcdavid".

Une base de données créée avec NLSCASE INSENSITIVE ignore le « case » uniquement pour les types de données NCHAR et NVARCHAR, mais traite les autres types de données caractères (CHAR, LVARCHAR, et VARCHAR) comme « case-sensitive ». Par ailleurs, il n'est pas possible d'utiliser dans une requête distribuée à la fois des bases de données sensibles et insensibles à la « case ».

Amélioration des requêtes utilisant la BTS datablade

Vous pouvez améliorer les résultats des requêtes utilisant la Basic Text Search Datablade en choisissant un analyseur de texte qui correspond le mieux à vos données et les besoins des requêtes. Un analyseur de texte détermine la façon dont le texte est indexé.

- L'analyseur « snowball » indexe de mots tiges.
- L'analyseur « CJK » gère les textes en chinois, japonais et coréen.
- L'analyseur « Soundex » indexe les mots par leur son.

D'autres analyseurs sont des variantes de ces analyseurs et l'analyseur standard. Vous pouvez également créer vos propres analyseurs.

Il est désormais possible de créer un thésaurus de synonymes.

Il est désormais possible de spécifier des « stopword » différents pour chacune des colonnes indexées au lieu d'utiliser le même « stopword » pour toutes les colonnes indexées.

Amélioration des performances des connexions réseaux

Vous pouvez améliorer les performances et la scalability des connexions réseau sur les systèmes UNIX en utilisant le paramètre de configuration NUMFDSERVERS. Utilisez ce paramètre de configuration pour ajuster le nombre maximal de poll threads à utiliser lors de la migration des connexions TCP / IP entre les processeurs virtuels (VPS).

Ajuster le nombre de poll threads, par exemple, en augmentant le nombre de poll threads au delà du nombre par défaut, est utile si le serveur de base subit un pic de nouvelles demandes de connection/déconnection ou si des contentions sont détectées par des verrous sur des Network Shared File (NSF).

Par exemple, si vous avez de multiples CPU VPs et poll threads et des NSF lock errors sont détectées, vous pouvez augmenter la valeur de NUMFDSERVERS pour réduire les contentions sur les NSF lock..

Informix Warehouse Accelerator

Informix Warehouse Accelerator est un produit Informix qui améliore les performances des requêtes analytiques de façon transparente à tous les utilisateurs d'Informix. L'accélérateur s'intègre de façon transparente dans un environnement Informix est basé une technologie où les données sont en mémoire.

Informix Warehouse Accelerator inclut un plugin d'interface graphique pour IBM Data Studio que vous pouvez utiliser pour effectuer les tâches d'administration de l'accélérateur. Vous pouvez utiliser l'accélérateur avec un serveur de bases de données utilisé pour le traitement transactionnel et décisionnel.

Retrouvez le Technical White Paper de Keshava Murthy en français sur <http://www.iiug.org/library/warehouse/>.

Technical Tip: Nouvelles fonctionnalités en 11.70.xC3

La version IBM Informix 11.70.xC3 sera disponible début juillet 2011. Voici en avant-première les nouvelles fonctionnalités:

Activation automatique du read-ahead

Vous pouvez configurer le serveur Informix afin d'utiliser automatiquement la lecture anticipée des pages (read-ahead) sur disque et ainsi améliorer les performances. Utilisez le paramètre de configuration `AUTO_READAHEAD` pour configurer automatiquement les opérations de read-ahead pour toutes les requêtes, et utiliser l'instruction `SET ENVIRONNEMENT AUTO_READAHEAD` pour configurer automatiquement les opérations de read-ahead pour une session particulière.

Le paramètre de configuration `RA_THRESHOLD` devient obsolète avec cette version.

Gestion de la mémoire

Il est désormais possible de configurer les actions que le serveur peut prendre afin de poursuivre son fonctionnement lorsque la mémoire est extrêmement faible. Vous pouvez spécifier des critères pour mettre fin à des sessions basées sur les temps d'inactivité (idle-time), l'utilisation mémoire, et d'autres facteurs de telle sorte que l'application ciblée puisse continuer à fonctionner et éviter des problèmes de out-of-memory. Cette fonctionnalité dite « low memory response » est utile pour les applications embarquées qui ont des limitations mémoire.

Par ailleurs, il sera possible d'activer dans le serveur la réservation d'une quantité spécifique de mémoire lorsque des activités critiques (telles que les activités de rollback) sont nécessaires ou quand le serveur est limité en mémoire disponible. Si vous activez le nouveau paramètre de configuration `LOW_MEMORY_RESERVE` en le fixant à une valeur en kilo-octets, les activités critiques pourront utiliser cette réserve mémoire et ainsi s'achever avec succès. Vous pouvez également ajuster dynamiquement la valeur du paramètre de configuration `LOW_MEMORY_RESERVE` avec les commandes onmode `-wm` ou `-wf`.

Amélioration du « Connection Manager »

Avant la version 3.70.xC3 vous deviez utiliser un Connection Manager différent pour gérer chaque type d'unité de connexion: Par exemple, vous deviez utiliser une instance du Connection Manager pour les serveurs de bases de données qui se trouvaient dans un GRID (ER) et une autre instance du Connection Manager pour les serveurs de base de données définis dans un cluster haute disponibilité (MACH11).

Vous pouvez configurer une instance unique de Connection Manager pour gérer les demandes de connexion du client vers un cluster de haute disponibilité (MACH11), un GRID, un Server Sets ou encore

vers des serveurs Entreprise Replication (ER)

Pour rendre la configuration plus simple, la plupart des options en ligne de commande sont obsolètes et les options sont définies en utilisant le fichier de configuration. De plus, le format du fichier de configuration est plus intuitif qu'auparavant.

Gestion du fichier de messages online.log

Vous pouvez utiliser des tâches du Scheduler pour réduire automatiquement la taille du fichier de messages online.log en tronquant ou en supprimant le journal des messages ou encore en configurant la rotation automatique du fichier.

De plus, vous pouvez utiliser les paramètres de la table ph_threshold pour spécifier le nombre maximum de fichiers de messages à conserver. Ces tâches et les paramètres associés sont utiles dans des environnements embarqués car ils réduisent les manipulations de l'administrateur dans la gestion des fichiers de log.

Ajout de built-in SQL fonctions

Le serveur base de données Informix intègre de nouvelles fonctions de manipulation de chaîne de caractères. Ces fonctions retournent soit une chaîne de caractères soit un entier qui décrit un string argument:

CHARINDEX(), INSTR(), LEFT(), LEN(), REVERSE(), RIGHT(), SPACE() et SUBSTRING_INDEX()

Cette version fournit également deux fonctions de type trigonométriques. Ces fonctions convertissent la valeur de l'argument dans les unités de mesure angulaire de radians en degrés, ou de degrés dans radians:

DEGREES(), RADIANS()

Conversion de code set avec Entreprise Replication

Entreprise Replication prend désormais en charge la réplication entre des serveurs de bases de données qui utilisent des code sets différents.

Vous pouvez convertir des serveurs vers le code set Unicode en minimisant l'indisponibilité de application, convertir des serveurs d'un code set à un autre, et répliquer les données entre des serveurs utilisant différentes locales.

Support des connexions shm et stream-pipes dans un Private Server.

Vous pouvez utiliser les connexions via la mémoire partagée et les stream-pipes entre le client et un serveur de bases de données qui a été installé sans les privilèges root (fonctionnalité Private Server introduite en version 11.50.xC2).

Dans le fichier sqlhosts, vous définissez la nouvelle option « cfg » pour spécifier un répertoire où seront stockés les fichiers nécessaires à la communication via shared memory et stream-pipe.

Gestion simplifiée des données de type Timeseries.

Gérer les données de type time series avec Informix TimeSeries devient plus facile que dans les versions précédentes:

- Stockage des données de TimeSeries: Containers pour les données TimeSeries sont créés automatiquement dans le même dbspace utilisé par la table. Les noms des containers et des sous-types TimeSeries peuvent aller jusqu'à 128 octets de long. Les tables et Containers TimeSeries peuvent utiliser des formats de page autres que le défaut.
- Création de TimeSeries: Vous pouvez utiliser des calendriers prédéfinis pour un time series. Vous pouvez désormais tester les données de type TimeSeries avec la base de données de démonstration stores_demo.
- Utilisation des données TimeSeries: Vous pouvez créer des tables virtuelles basées sur le résultat d'une expression TimeSeries et vous pouvez utiliser les virtual tables pour mettre à jour les données. Vous pouvez également extraire des données TimeSeries au format XML avec la fonction TSToXML().
- Suppression des données TimeSeries: Vous pouvez supprimer rapidement de grandes quantités de données de séries chronologiques.

L'IBM Informix TimeSeries Plug-in pour l'outil OpenAdmin (OAT) fournit une interface graphique pour l'examen et à l'administration des données TimeSeries.

Etude de cas : Objectif performance...

"Bonjour, merci de vous être connecté sur notre station. Votre mission est, si vous l'acceptez, d'implémenter une nouvelle table dans la base de données, d'en faire un chargement initial avec un temps d'opération inférieur à 10 minutes, et de fournir à l'équipe de développeurs une manière de gérer au mieux les temps de réponses sur une requête imposée qui, selon le comité des testeurs de l'application reportant au Directeur des Opérations, a un temps de réponse inacceptable.

Vous disposez d'IBM Informix Dynamic Server V11.70 FC2 Express Edition, un serveur Linux avec un processeur 4core, 16 Gb de RAM entièrement à votre disposition.

Ah! J'oubliais, la table à charger comporte environ 70 millions de rangées, mais n'est pas très complexe quant à son schéma et son indexation. Il faudra de plus y charger environ 500 000 rangées tous les soirs"

Ce genre de demande vous rappellera très certainement des situations vécues soit en tant que DBA, soit en tant que développeur. On constate généralement, à ces instants précis, un certain flottement dans le regard du destinataire de cette requête, de même que des symptômes caractéristiques du découragement ou d'une féroce envie de décliner la demande sous le chef de « ça n'est pas réaliste ». Il est vrai que dans les versions antérieures d'IDS cette demande représentait un véritable défi, difficile à relever. Mis à part peut-être l'utilisation de High Performance Loader, pas trop simple à mettre en œuvre dans l'urgence, il faut avouer qu'il y avait au minimum de quoi être troublé par cette demande.

Fort heureusement, nous sommes en 2011 et les laboratoires IBM Informix nous ont concocté un certain nombre d'améliorations très utiles, qui vont nous permettre de relever le défi avec IDS 11.70.

Analysons tout d'abord la demande:

1) La nouvelle table contient tous les billets d'avion sur les liaisons entre la France, l'Allemagne, l'Angleterre, l'Italie, l'Espagne et les Etats-Unis depuis le 1er janvier 2005. en voici le schéma :

```
create table tickets (
    ticketno bigint,           # N° unique de ticket
    lastname varchar(60),     # Nom voyageur
    firstname varchar(30),    # Prénom voyageur
    gender char(3),           # MR/MRS/INF
    flightdate date,          # Date du vol
    flightno char(6),         # N° du vol
    fileno char(8),           # N° de dossier
    price money(8,2),         # Prix du vol
    carrier char(2),          # Compagnie aérienne
    departfrom char(3),       # aéroport de départ
    destination char(3)       # aéroport de destination
);

create index tickets3 on tickets (flightdate);
create index tickets7 on tickets (destination);
```


Et enfin la requête qui pose des problèmes à nos amis développeurs:

```
select destination , count(*)
from tickets,airports
where tickets.destination = airports.code
and tickets.destination between "AAA" and "ZZZ"
and airports.country = 'France'
and flightdate between "21/03/2006" and "31/07/2010"
group by 1
```

J'avais au préalable réussi à convaincre les développeurs de fournir des indexes mono-colonne, de façon à éviter l'obtention de trop nombreux niveaux de B-Tree affaiblissant la performance. L'évocation du chiffre de 70 millions de rangées initiales, grossissant environ d'un million tous les 2 jours, est par contre de nature à susciter au minimum de l'intérêt, au maximum de l'inquiétude quant à l'atteinte des objectifs définis.

Nous ne parlerons pas aujourd'hui de fragmentation de cette table qui pourrait nous apporter une solution ultérieurement en cas de détection de problèmes d'IO sur celle-ci. Nous ne parlerons pas non plus de la compression des tables, technique provenant de XPS et implémentée à partir de la version 11.50.xC6 qui, tout en réduisant environ de moitié (ou plus) l'espace disque, permet d'augmenter de façon significative les performances sur les "grosses" tables. Cette fonctionnalité est pour l'instant « optionnelle » (lire payante), et ne fait donc pas partie des options à envisager dans notre cas.

Concentrons nous donc sur la première partie du défi : le chargement de 70 millions de rangées en moins de 10 minutes.

Nous avons pensé à High Performance Loader, qui est un outil très efficace, mais un peu complexe à mettre en oeuvre pour une première fois et dans le temps imparti. Je voulais donc évoquer ici l'utilisation des tables externes, provenant également de XPS et implémentées depuis la version 10.50.

Le principe est très simple, vous créez votre table, sauf qu'au lieu de résider dans l'instance IDS, celle-ci est constituée par un fichier sur file system, structuré soit par l'existence de délimiteurs de champs, soit par une structure fixe. Vous créez la table externe qui pointe sur ce(s) fichier(s) et qui en décrit la structure comme une table « normale ». Nous avons donc appliqué ceci pour le cas d'aujourd'hui :

```
CREATE EXTERNAL TABLE ext_tickets SAMEAS tickets
  USING (DATAFILES ("DISK:/home/eric/StagingArea/ObjectifPerf/tickets.unl"),
  DELIMITER "|",EXPRESS,MAXERRORS 10,REJECTFILE "/tmp/x_tickets.out")
```

Expliquons :

- « SAMEAS tickets » veut dire créer la table à l'image du schéma de la table déjà existante tickets.
- « DATAFILES DISK » : notre fichier est sur disque et se nomme...
- « DELIMITER » : c'est une structure avec délimiteurs de champs, ici le « pipe »
- « EXPRESS » : le mode de chargement. Il s'oppose au mode DELUXE. Ce mode évite l'utilisation des pages de buffers lors de la lecture dans cette table, ce qui va muscler fortement la performance de notre chargement.
- « MAXERRORS » : au-delà de ce nombre d'erreurs, la création de la table s'arrête.
- « REJECTFILE » : le fichier où sont consignées les erreurs de chargement.

Une table externe peut donc être manipulée par des tous types de requêtes SELECT (order by, group by, prédicats) , mais ne peut pas être UPDATED,DELETED ni INSERTED.

Résultat , avec table raw

la commande de chargement est :

```
INSERT INTO tickets SELECT * FROM ext_tickets
70000000 row(s) inserted.
real 6m20.937s
user 0m0.005s
sys 0m0.005s
```

Nous avons comparé avec la méthode `insert into table load from file`, et le résultat est sans appel : nous avons arrêté le chargement au bout de 24 minutes : seulement 24 Millions de rangées étaient chargées.

Autre point important à relever : nous avons créé une table de type « RAW ». Ce type de table a la particularité de désactiver la journalisation de la table incriminée. En conséquence, la performance de chargement est drastiquement améliorée du fait de l'absence d'IO à cause de la journalisation (logical logs). La différence est plus que notable. Le gros intérêt est qu'il est possible de changer en type STANDARD (journalisation) après le chargement par un simple ALTER TABLE, puis repasser en RAW pour les chargements consécutifs. Nous évitons également l'écueil de la 'long transaction'.

Nous avons donc atteint le premier objectif de passer sous les 10mn de temps de chargement, et restons sereins pour les chargements journaliers.

Passons maintenant au deuxième défi : la requête

La table tickets a un nombre de rangées respectable (70 millions de rangées), et à cette échelle il convient de veiller à optimiser les ressources ainsi que les accès. Nous partons de la baseline proposée par l'équipe des développeurs et prendre une mesure de performance.

Nous avons de fait renoncé à la baseline, comprenant les plaintes des développeurs : la requête a été abortée au bout de 60 mn sans qu'elle soit terminée, voici son plan de requête :

```
Estimated Cost: 3126460
Estimated # of Rows Returned: 775663
Temporary Files Required For: Group By

1) eric.airports: INDEX PATH
   Filters: (eric.airports.code >= 'AAA' AND eric.airports.code <= 'ZZZ' )
   (1) Index Name: eric.i_airp3
       Index Keys: country (Serial, fragments: ALL)
       Lower Index Filter: eric.airports.country = 'France'

2) eric.tickets: INDEX PATH
   Filters: (eric.tickets.flightdate >= 21/03/2006 AND
            eric.tickets.flightdate <= 31/07/2010 )
   (1) Index Name: eric.tickets7
       Index Keys: destination (Serial, fragments: ALL)
       Lower Index Filter: eric.tickets.destination = eric.airports.code

NESTED LOOP JOIN
```

Nous obtenons donc une nested loop (boucle imbriquée pour la jointure), avec utilisation de l'index sur les pays pour la table airports (10000 rangées), et l'index date du vol pour la table des tickets. Onstat - p nous donne 7509528 rangées lues (read).

C'est à ce moment que nous introduisons la nouvelle directive d'optimiseur MULTI_INDEX, qui va nous permettre d'utiliser plusieurs indexes sur une même table, contrairement à ce qui se faisait auparavant.

La syntaxe :

```
SELECT --+multi_index(tickets tickets3 tickets7)
distinct(destination) , count(*) ....
from tickets,airports etc....
```

Nous allons dire à l'optimiseur d'utiliser les index tickets3 **ET** tickets7 pour la table tickets afin de lire un minimum de pages de données, et par conséquent réduire les IO sur cette requête. Voici le résultat du set explain :

DIRECTIVES FOLLOWED:
MULTI_INDEX (tickets tickets3 tickets7)
DIRECTIVES NOT FOLLOWED:
Estimated Cost: 113176408
Estimated # of Rows Returned: 775663
Temporary Files Required For: Group By

1) eric.tickets: MULTI INDEX PATH (SKIP SCAN)

(1) Index Name: eric.tickets7
Index Keys: destination (Serial, fragments: ALL)
Lower Index Filter: eric.tickets.destination >= 'AAA'
Upper Index Filter: eric.tickets.destination <= 'ZZZ'

AND

(2) Index Name: eric.tickets3
Index Keys: flightdate (Serial, fragments: ALL)
Lower Index Filter: eric.tickets.flightdate >= 21/03/2006
Upper Index Filter: eric.tickets.flightdate <= 31/07/2010

2) eric.airports: INDEX PATH

Filters:

Table Scan Filters: (eric.airports.code >= 'AAA' AND eric.airports.code <= 'ZZZ')

(1) Index Name: eric.i_airp3
Index Keys: country (Serial, fragments: ALL)
Lower Index Filter: eric.airports.country = 'France'

DYNAMIC HASH JOIN

Dynamic Hash Filters: eric.tickets.destination = eric.airports.code

La physionomie du plan de requête a totalement changé : nous voyons l'activation de la directive multi-index sur la table tickets, qui a pour effet l'utilisation de l'index tickets7 + l'index tickets3 et finalisant le plan par un HASH JOIN pour la jointure.

Onstat -p nous indique que seulement 2.294.947 de rangées ont été lues, ce qui devrait signifier un progrès sensible. Et le temps de réponse ????

Le voici :

```
155 row(s) unloaded
real  5m22.672s
user  0m0.005s
sys   0m0.008s
```

Oui ! c'est bien 5 mn 22s. L'objectif d'amélioration du temps de réponse est atteint. N'ayant pas de données implémentées sur plusieurs disques, nous ne pourrons pas profiter du multiplexage des scan threads qui auraient pu améliorer encore notre performance. N'ayant pas lancé ce test sur la Ultimate Edition, nous ne pouvons pas bénéficier du PDQPRIORITY qui aurait sans doute amélioré légèrement les résultats.

Nous avons clairement profité de la directive MULTI INDEX, qui a diminué le temps d'attente de plus d'une heure (peut-être beaucoup plus) à 5 mn 22s. Il faut savoir également que cette technique est extrêmement efficace dans des requêtes avec prédicat « OR » portant sur des indexes différents d'une même table. Combien de fois nous a-t-il fallu réécrire une requête de ce type en incluant une UNION ? Avec cette technique, plus la peine de tout réécrire, il suffit d'ajouter la directive d'optimiseur qu'il faut, et le tour est joué !

Nous vous encourageons également à utiliser l'option de onmode '-Y sessionid 1 output file', qui produira le fichier d'explain de la session 'sessionid' sans avoir à rajouter le 'SET EXPLAIN ON' dans le code de l'application. A utiliser toutefois avec modération en environnement production, car ce mode est très volubile et peut générer une charge d'IO qui pourra parasiter la performance.

Voilà, mission accomplie ! Merci aux techniques de tables externes en mode EXPRESS, de table de type RAW qui suppriment temporairement la journalisation, et la nouvelle directive d'optimiseur MULTI_INDEX.

A bientôt sur notre station pour de nouveaux défis.

Derniers articles

- Enhanced availability and workload management with Informix flexible grid: [voir l'article](#)
- An insider's guide to star join optimization in IBM Informix Server: [voir l'article](#)
- Informix DBA: Fastest Informix DBA contest III: [voir l'article](#).
- IIUG user view: Big data, big time: [voir l'article](#)
- Take advantage of fragment-level statistics and smarter statistics in IBM Informix: [voir l'article](#)
- Tune your Informix database for top performance, Part 1: [voir l'article](#)
- IBM Redbook: Migrating from Oracle to IBM Informix Dynamic Server: [voir l'article](#).

Vidéos

- Advanced Data Tools benefits from integrated IBM Informix and BladeCenter solution: <http://www.youtube.com/watch?v=dW3Y70moMoc>
- IBM Informix helping create smarter planet, smarter retail: http://www.youtube.com/watch?v=D1LbxxfD_tY
- Informix users sharing the business value of Informix! <http://www.youtube.com/watch?v=aYaP1x68JoU>

Liens Utiles

- Informix Developer Works : <http://www.ibm.com/developerworks/data/products/informix/>
- IBM Redbook : <http://www.redbooks.ibm.com/portals/data>
- IBM Data Management magazine: <http://www.ibm.com/developerworks/data/dmmag/>
- IIUG : <http://www.iiug.org/index.php>
- IBM et Virtualisation : <http://www-03.ibm.com/systems/virtualization/>
- Informix et le BETA program : <http://www-01.ibm.com/software/data/informix/beta/>
- Informix sur facebook: <http://www.facebook.com/IBM.Informix>
- Informix sur Twitter : http://twitter.com/IBM_Informix
- The IIUG forums: <http://www.iiug.org/forums/technical.php>
- Blogs, Videos, News and more at <http://planetids.com>

Informix blogs

- Bruce Weed's blog <http://bruceweed.wordpress.com/>
- Fernando Nunes: Informix Technology: <http://informix-technology.blogspot.com/>
- Eric Vercelletto : Le village Informix <http://levillageinformix.blogspot.com/>
- Jacques Roy: <http://www.ibm.com/developerworks/blogs/page/jacquesroy>
- Informix blogs: <http://www.ibmdatabasemag.com/blog/main/archives/informix/index.html>
- The Informix Zone: <http://www.informix-zone.com>
- The Informix mag: <http://www.informixmag.com/>

Abonnement / Annulation / Avis

Cette newsletter est envoyée à des adresses enregistrées. Si vous souhaitez respectivement vous abonner ou vous désabonner, veuillez envoyer un mail avec pour sujet « ABONNER » ou « DESABONNER » à l'adresse email : ifmxnewsletter@fr.ibm.com.

Votre avis et vos contributions sont bien entendu les bienvenus ! N'hésitez pas à nous les faire parvenir à l'adresse email : ifmxnewsletter@fr.ibm.com.

Les contributeurs de ce numéro

Khaled Bentebal	Président du User Group Informix France Directeur Général – Consultix
Eric Vercelletto	Directeur Général – BeGooden ITConsulting
Olivier Bourdin	EMEA Informix L3 Advanced Problem Diagnostic IBM Certified Products Services IBM Software Group, Information Management
Yoram Benchetrit	EMEA Informix L3 Advanced Problem Diagnostic IBM Software Group, Information Management
Frédéric Delest	EMEA Informix L2 Diagnostic and System Down IBM Software Group, Information Management