TechTarget

Maitrisez (mieux) SQL Server

Les fonctions essentielles, les focntionnalités BI, les bonnes pratiques de sécurité

Présentation

Système de gestion de bases de données de Microsoft, SQL Server, aujourd'hui en version 2014, représente la vision de l'éditeur en matière de gestion des données.

La base, 3e du marché dans le monde, est au coude à coude avec IBM - toutes deux avoisinent les 20% de parts de marché - mais se positionne encore loin derrière Oracle.

Pour combler son retard, Microsoft SQL Server s'est habillée, au fil des versions, aux couleurs des tendances des dernières saisons.

La dernière version en date se dote, par exemple, d'une technologie de In-Memory ou encore d'une option de stockage en colonnes pour se rapprocher des technologies dites NoSQL.

Autre avancée, la solution a amorcé un virage très BI.

Quant à la sécurité, Microsoft ne cesse de répéter qu'il l'améliore.

Mais encore faut-il maîtriser toutes ces fonctionnalités d'un SGBD de plus en plus mature.

C'est ce que vous propose de faire de ce Guide Essentiel en abordant, concrètement, des fonctions clefs à connaître absolument, des fonctions orientées BI et des bonnes pratiques de sécurité que beaucoup connaissent certainement, mais que beaucoup oublient aussi (trop) souvent ■

La Rédaction

TechTarget / LeMagIT

SQL Server 2014 : 8 fonctions à retenir

SQL Server 2014, lancée début avril 2014. Voici une liste des huit nouvelles fonctionnalités clés embarquées dans la dernière version du SGDB de Microsoft.

Un moteur OLTP en mémoire renforcé

Le moteur OLTP (Online Transaction Processing) In-Memory, nom de code Hekaton, permet de créer des tables en mémoire optimisées au sein d'une base de données relationnelles traditionnelle.

Ce moteur résout par exemple des problèmes liés à des situations de requêtes hautement concurrentes, car il s'adosse à des structures de données dépourvues de verrou (latches). Ce qui signifie qu'il n'y a aucun verrou logique sur les chemins critiques et donc pas d'impact sur les performances au sein du système.

En revanche, le moteur utilise une technique MVCC (multi-version concurrence control) optimiste qui fournit des niveaux d'isolement des transactions - ce qui contribue à éviter les interférences entre transactions. Ainsi tous les processus utilisateurs peuvent accéder à chaque ligne d'une table sans verrous logiques ni latches.

L'association de ce MVCC et de cette structure de données donne un système au sein duquel les processus utilisateurs peuvent s'exécuter sans attente ni blocage. De plus, les procédures stockées, qui sont exécutées dans des tables optimisées en mémoire, via Transact-SQL, sont en fait compilées en un code machine efficace.

Cela optimise les performances du runtime pour certains workloads et certains types de requêtes, car le code machine généré ne contient que ce qui est nécessaire pour effectuer la requête, et rien de plus. Selon Microsoft, certaines applications peuvent atteindre des gains des performances d'un facteur de 50, simplement en ayant recours à ce moteur OLTP en mémoire.

Cette fonction est seulement supportée par les éditions Enterprise 64-bit, Developer ou Evaluations de SQL Server 2014. Pour utiliser ce moteur, qui est sans conteste la plus importante nouveauté du SGBD, vous pouvez consulter les ressources publiées sur le site TechNet de Microsoft.

Vous pouvez également utiliser l'assistant SQL Server 2014 Memory Optimization Advisor qui peut être lancé à partir de SQL Server Management Studio (SSMS) pour

identifier et migrer les tables compatibles en mémoire, puis choisir les procédures stockées qui peuvent être compilées dans du code machine pour une exécution plus performante.

Des index de données en colonnes In-Memory améliorés

Dans SQL Server 2012, Microsoft avait présenté un nouveau type d'index non clusterisés, des index de données en colonnes en mémoire xVelocity. Le problème est que est que ces index ne pouvaient pas être mis à jour dans SQL Server 2012 : ainsi une fois l'index créé, nous ne pouviez pas directement ajouter, supprimer et ou modifier les données dans la table sous-jacente. Pour mettre à jour cette dernière, l'index devait d'abord être désactivé, puis recréé une fois les données à jour. SQL Server 2014 améliore ces index en colonnes. Dans la dernière mouture du gestionnaire de bases de données, ceux-ci sont disponibles à la fois en mode clusterisés et non clusterisés, et peuvent être mis à jour.

Extension du pool tampon in-memory

Cette nouvelle capacité donne à chaque noeud SQL Server la possibilité de disposer de son propre SSD, un type de mémoire vive non volatile, en guise de buffer. La configuration serveur permet à un pool de mémoire tampon d'être pris en compte par les workloads OLTP. Cela contribue à résoudre les problèmes de goulet d'étranglement des I/O tout en améliorant la bande passante I/O, du fait des performances et de la faible latence des SSD.Le fait que la Flash soit un stockage non volatile permet aussi de se protéger contre les risques de pertes de données.

Voici la commande pour activer cette fonction :

```
ALTER SERVER CONFIGURATION
 SET BUFFER POOL EXTENSION
 { ON ( FILENAME = 'os_file_path_and_name'
     ,SIZE = <size> [ KB | MB | GB ] )
 | OFF }
```

Cette fonction est seulement supportée par les versions Enterprise 64-bit, Developer et Evaluation de SQL Server 2014.

Gouverneur de ressources

SQL Server 2014 apporte aussi deux nouveaux paramètres pour le gouverneur de ressources, qui peuvent être utilisés pour contrôler les I/O physiques allouées à des threads pour un pool de ressources donné. Ces

nouveaux paramètres sont : MIN_IOPS_PER_VOLUME et MAX_IOPS_PER_VOLUME - et fixent le minimum et le maximum d'opérations par seconde et par disque pour un pool de ressource donné. Pour plus d'informations, consultez ALTER RESOURCE GOVERNOR (Transact-SQL).

Amélioration de la disponibilité du SGBD

SQL Server 2014 intègre également la possibilité de constituer des groupes de disponibilité avec des VM Azure. Ce qui signifie que nous pouvons désormais utiliser des VM Azure comme répliques (ou replicas, en langage Microsoft) pour des groupes de disponibilité asynchrones. De plus, SQL Server 2014 supporte désormais jusqu'à 8 réplicas secondaires disponibles en lecture, même lorsqu'ils sont déconnectés du réplica primaire. Les groupes de disponibilité AlwaysOn supportent également les fonctions OLTP en mémoire.

Sauvegarde intégrée à Azure

SSMS dans SQL Server 2014 vous permet de sauvegarder directement vos bases de données existantes sur site dans le service de stockage en cloud Windows Azure Storage. Vous pouvez également restaurer des bases de données à partir d'Azure.

Amélioration de la sécurité

Microsoft SQL Server 2014 apporte également de nouvelles permissions pour la gestion de la sécurité. Ces nouvelles permissions serveurs sont : CONNECT ANY DATABASE, IMPERSONATE ANY LOGIN et SELECT ALL USER SECURABLES. Elles permettent d'attribuer des permissions aux administrateurs de base de données pour qu'ils puissent s'acquitter seulement de leurs tâches et sans pouvoir accéder aux données utilisateurs.

Chiffrement des sauvegardes

SQL Server 2014 donne enfin la possibilité de chiffrer les données stockées dans la base de données lors d'une opération de sauvegarde.

Business Intelligence : 5 fonctions de SQL Server 2012 à connaître

SQL Server 2012 propose certaines nouvelles fonctions qui améliorent et étendent ses fonctionnalités bien au-delà de SQL Server 20008 et 2008 R2. Et cela est particulièrement vrai dans le domaine de la Business Intelligence (BI), où les améliorations en matière de reporting et de capacités analytiques sont plus marquées. Cinq de ces fonctions méritent que les entreprises s'y attardent.

Modèle sémantique BI

SQL Server 2012 embarque un modèle sémantique BI (BI Semantic Model - BISM) afin de mettre à disposition un framework de conception pour doter certaines plates-formes analytiques et de reporting de fonctions de BI. Bien que BISM ne soit finalement pas un produit à part entière, il fournit la structure pour créer des modèles physiques dans SQL Server Analysis Services (SSAS) et PowerPivot pour Excel. Dans SSAS, vous pouvez créer deux types de modèles BISM : multidimensionnel ou tabulaire. Dans PowerPivot pour Excel, vous ne pouvez créer que des modèles tabulaires.

Le modèle multidimensionnel offre une approche somme toute traditionnelle. Il est aligné sur Unified Dimensional Model (UDF) des versions précédentes de SSAS, avec les données organisées en cube et dimensions. Comme avec UDF, le modèle multidimensionnel s'appuie sur le langage Multidimensional Expressions (MDX) pour interagir avec les datastores multidimensionnels, proposant ainsi un environnement puissant pour effectuer des opérations analytiques complexes.

Dans SQL Server 2012 se distingue le modèle tabulaire, qui à l'inverse du modèle multidimensionnel, organise les données dans des tables, avec des lignes et des colonnes, comme vous pouvez le trouver dans les bases de données relationnelles. De plus, ce modèle utilise le langage Data Analysis Expressions (DAX) pour accéder aux données et peut en extraire à partir d'un grand nombre de sources, telles que les SGBD, les cubes SSAS, les fichiers texte et les workbooks PowerPivot. Le modèle tabulaire a été lancé avec SQL Server 2008 R2, lors de la sortie de PowerPivot pour Excel et PowerPivot pour SharePoint. Le modèle s'adosse au moteur xVelocity (anciennement VertiPaq) pour cacher les données en mémoire tout mettant à disposition des algorithmes de compression et de scanning pour supporter les analyses de données haute-performance.

Index columnstore

Cette technologie xVelocity nous amène à la 2^{e} des 5 fonctions de BI de SQL Server 2012 : les index

columnstore non clusterisés. Comme les technologies de ce type, un index columnestore est défini sur une ou plusieurs colonnes de tables. A l'inverse des indexes non clusterisés, la donnée existe dans un format en colonne. Autrement dit, la donnée dans chaque colonne indexée est stockée dans sa propre colonne à l'intérieur de l'index. Si une table contient plus d'un million de lignes, l'index est compartimenté en segments, mais la structure en colonne est conservée. Les index non clusterisés stockent les données en lignes.

En plus de ce stockage en colonnes, les index dits columnstore utilisent les fonctions de stockage et de compression avancées de xVelocity. Ces fonctions peuvent améliorer les performances des requêtes effectuées pour extraire ou traiter de grand volume de données , ce type de requête généralement associé à l'entrepôt de données – où les données sont souvent groupées, filtrées, agrégées puis reliées à travers plusieurs tables.

Les gains de performances sont souvent attribués au fait que seules les bonnes colonnes sont montées en mémoire, les données en mémoire sont compressées plus efficacement, et les requêtes sont optimisées pour des traitements analytiques.

Data Quality Services

Pour que le reporting, l'analyse et l'entreposage soient efficaces, il faut que la donnée soit elle-même correcte. Mais extraire des données de plusieurs systèmes, dans un standard différent, est généralement source d'erreurs, de corruptions de données et d'incohérences. C'est justement là que Data Quality Services (DQS) entre en jeu. Nouveauté dans SQL Server 2012, DQS donne les outils adéquats pour résoudre les problèmes liés à l'inexactitude des données, leur incohérence et leur duplication. Les gestionnaires des données (data steward) ainsi que les professionnels de l' IT peuvent ainsi utiliser DQS pour nettoyer les données et s'assurer qu'elles sont adaptées à leur outil de BI ainsi qu'à leurs besoins métiers.

L'environnement DQS comporte deux éléments : le Data Quality Server et le Data Quality Client. Le serveur prend en charge toutes les opérations lourdes. Il héberge le moteur DQS, stocke les informations liées au projet et gère les bases de connaissances.

Une base de connaissance est un dossier au sein duquel l'information, ou la connaissance, est déposée. Elle identifie les potentielles inexactitudes des données et propose des mesures correctives.

La base de connaissances comprend un ou plusieurs domaines, chacun contenant les connaissances pour un type spécifique de données. Par exemple, un domaine peut contenir les connaissances nécessaires pour s'assurer que les provinces canadiennes soient toutes référencées de la même façon et que seules les provinces correctement formées soient inclues.

Le Data Quality Client propose une interface pour administrer DQS, gérer les bases de connaissances et contrôler les projets liés à la qualité des données qui font correspondre les connaissances aux données.

Power View

Autre nouvelle fonction de BI : Power View, un add-in SQL Server Reporting Services pour SharePoint Server. Power View s'appuie sur le framework Silverlight et propose aux utilisateurs des outils Web pour explorer les données et créer des rapports ad hoc à base de méthodes de visualisation de données enrichies. Créer un rapport Power View est identique à créer un tableau croisé dynamique dans Excel. Les utilisateurs travaillent toujours avec leurs données en cours et n'ont jamais à passer d'une vue à une autre, à l'inverse de Report Builder et Report Designer dans SQL Server Data Tools.

Les utilisateurs lancent Power View d'un site SharePoint configuré avec l'add-in SSRS. SharePoint Server et SQL Server doit donc tous deux être installés. Un développeur ou un administrateur doit configurer au moins un modèle de donnée. Celui-ci sert d'interface entre les rapports et les sources de données, facilitant ainsi l'accès aux données requises, sans avoir à comprendre la structure de la donnée sous-jacente.

Pour créer des rapports, les utilisateurs doivent seulement s'identifier dans le site SharePoint (via un navigateur qui supporte Silverlight), localiser le bon modèle de donnée dans une bibliothèque de documents ou une gallerie PowerPivot et lancer Power View depuis ce modèle. Ils peuvent ensuite créer leurs rapports via de simples glisser-déposer.

Recherche sémantique

SQL Server a certes ses racines dans le monde des données structurées et relationnelles, mais plus que jamais, la BI comprend également des données non structurées dans ses traitements analytiques et ses rapports. SQL Server répond à cela et s'accommode du flux entrant de données non structurées en fournissant des fonctions de recherche plein texte, intégrées à la base de données.

La recherche plein texte permet aux requêtes d'opérer des recherches sur les mots et non pas les données dans des colonnes, y compris les colonnes FileStream qui pointent sur les fichiers de données non structurées. Toutefois, SQL Server 2012 va plus loin avec la recherche sémantique, une fonction qui étend la recherche plein texte pour extraire et indexer de façon statistique les phrases clés des documents non structurés.

La recherche sémantique va plus loin que la recherche plein texte en s'intéressant au sens du document plutôt que seulement à ses mots. Cela rend possible des fonctions, comme l'extraction automatique de tag, la découverte de contenus associés ou la navigation hiérarchique entre des contenus sur un sujet identique. La recherche sémantique aide à trouver des documents qui sont identiques ou liés à ceux recherchés.

Une autre fonction de SQL Server 2012 rend cette recherche sémantique encore plus puissante : Filetable. Une FileTable est un type de table qui étend les fonctions de FileSteam pour supporter les API Win32. Chaque ligne représente un fichier ou un dossier et est accessible directement depuis une application Windows. Dans l'application, les fichiers et les dossiers apparaissent comme s'ils étaient stockés dans le système de fichiers, plutôt que dans la base de données. Associée à FileTable, la recherche sémantique devient un élément clé pour supporter des applications qui utilisent des données non structurées et extraire de l'information pertinente.

BI et SQL Server 2012

Ces 5 fonctions de BI de SQL Server peuvent s'avérer être des outils efficaces lorsque vous décidez de déployer une plate-forme de BI. Toutefois, ce qui est couvert ici ne représente que certaines des fonctions qui ont été mises à jour ou ajoutées dans SQL Server 2012. Par exemple, SSAS propose de nouvelles fonctions DAX et l'usage des ressources en matière de reporting pour les bases multidimensionnelles a été amélioré. SQL Server Integration Services, ou SSIS, comprend désormais un outil DQS Cleansing et a amélioré sa gestion de la mémoire des transformations Merge et Merge Join. L'intégration de SQL Server SSRS à SharePoint a été revue pour mieux exploiter certaines fonctions de SharePoint.

*– **Robert Sheldon**, consultant technique et auteur de nombreux ouvrages et articles sur Windows, sur les SGBD et sur l'implémentation d'outils de BI*

SQL Server : 8 bonnes pratiques de sécurité

Parmi les responsabilités clés de l'administrateur de bases de données, celui-ci doit s'assurer que toutes les instances SQL Server dont il s'occupe sont sécurisées. La sécurité de SQL Server est en soi un sujet particulièrement vaste. Cet article n'est qu'une introduction aux huit bonnes pratiques que je vous propose de suivre pour sécuriser les instances SQL Server que vous administrez.

1 - De l'importance des sauvegardes de bases de données cryptées

Pour l'entreprise, les sauvegardes de bases de données revêtent toujours une importance vitale. Si les fichiers de sauvegarde ne sont pas cryptés, ils sont faciles à copier et à restaurer sur n'importe quelle autre installation SQL Server ; une situation qui favorise le vol des données et amoindrit la sécurité. Pour éviter ce scénario peu engageant, l'administrateur de bases de données peut créer des sauvegardes en utilisant la fonction intégrée MEDIAPASSWORD. L'exemple de script ci-dessous permet de créer des sauvegardes de bases de données cryptées dans SQL Server :

```
BACKUP DATABASE AdventureWorks
TO DISK='C:\AdventureWorks.BAK'
WITH MEDIAPASSWORD='C0mplexP@ssW0rd'
GO
```

2 - Sécuriser le dossier de sauvegarde de base de données en éliminant les utilisateurs superflus

Un administrateur de bases de données doit s'assurer que l'accès au dossier de sauvegarde est restreint, qu'il n'est accordé qu'aux utilisateurs qui en ont vraiment besoin. Un accès non autorisé peut rendre ce dossier vulnérable, les fichiers de la sauvegarde pouvant alors être copiés sur des serveurs distants. Parallèlement, ce type d'accès peut favoriser la suppression accidentelle de fichiers de sauvegarde vitaux ; suppression qui détruirait la séquence de restauration de la base de données. Aussi est-il essentiel que l'administrateur de bases de données s'assure que seules les personnes appropriées disposent d'un accès au dossier de sauvegarde de la base de données.

3 - Utiliser l'authentification Windows plutôt que le mode d'authentification SQL Server

L'usage de l'authentification Windows pour se connecter à SQL Server constitue une bonne pratique de sécurité. Dans SQL Server, le mode d'authentification Windows permet d'exploiter les politiques en vigueur à l'échelle de l'entreprise concernant Active Directory, les comptes, les groupes et les mots de passe. L'accès devient ainsi plus sécurisé.

Si vous utilisez le mode d'authentification SQL Server pour vous connecter, il est déconseillé d'utiliser un compte d'administrateur système (SA). Utiliser une authentification SQL Server ne vous empêche pas de tirer parti des politiques de mot de passe de Windows lorsque vous définissez les mots de passe des comptes d'utilisateur.

4 - Complexifier le mot de passe du compte de l'administrateur système

Si vous utilisez une authentification en mode mixte dans SQL Server, définissez systématiquement un mot de passe complexe pour le compte SA. Une bonne pratique consiste à toujours éviter de connecter des applications Web à SQL Server en utilisant le compte SA.

Il faut toujours éviter d'utiliser le compte de l'administrateur système pour procéder à des activités de maintenance au quotidien.

Pour les tâches quotidiennes, utilisez des comptes Windows dotés des autorisations adéquates. Bonne pratique toujours, pensez à changer les mots de passe des comptes SA au bout de plusieurs jours.

5 - Auditer les connexions

Dans le cadre de vos bonnes pratiques de sécurité SQL Server, auditez systématiquement les échecs de connexion à votre SGBDR.

Une fois l'audit des connexions activé dans SQL Server, les informations relatives aux connexions en échec ou réussies sont consignées dans les journaux des erreurs.

Une surveillance régulière de ces journaux permet d'identifier ponctuellement des activités suspicieuses.

6 - Désactiver le service SQL Server Browser

Autre bonne pratique de sécurité SQL Server, l'administrateur de bases de données doit désactiver le service SQL Server Browser lorsqu'il exécute une instance par défaut de SQL Server.

Même si vous exécutez une instance nommée, vous pouvez, pour vous y connecter, définir explicitement le port et le mentionner au sein des chaînes de connexion des applications concernées.

7 - Désactiver les fonctions inutilisées dans SQL Server

Lorsqu'elles sont inusitées, désactivez les fonctions telles que XP_CMDSHELL, OLE AUTOMATION, OPENROWSET et OPENDATASET pour réduire la surface d'exposition aux attaques. Le gestionnaire de configuration de Microsoft SQL Server 2005 permet d'activer et de désactiver ces fonctions.

Si vous utilisez SQL Server 2008 ou une version ultérieure, vous ferez appel pour cela à la fonction de gestion basée sur des stratégies.

8 - Diminuer les privilèges du compte des services SQL Server

Enfin, dernière bonne pratique, un administrateur de bases de données doit toujours exécuter les services SQL Server au moyen d'un compte de domaine local doté de privilèges minimaux.

Evitez d'exécuter les services SQL Server dans le cadre de comptes de système local, d'administrateur local ou d'administrateur de domaine.

Toutefois, assurez-vous que le compte des services SQL Server dispose d'une autorisation « Contrôle total » en lecture et en écriture sur les répertoires des données, des journaux et des sauvegardes.

– ***Ashish Kumar Mehta***

AUTHEURS

Le document consulté provient du site www.lemagit.fr

Cyrille Chausson | *Rédacteur en Chef*

Philippe Ducellier | *Journaliste*

Linda Koury | *Directeur Artistique*

Neva Maniscalco | *Designer*

TechTarget
22 rue Léon Jouhaux, 75010 Paris
www.techtarget.com

TechTarget édite des publications pour les professionnels de l'IT. Plus de 100 sites qui proposent un accès rapide à un stock important d'informations, de conseils, d'analyses concernant les technologies, les produits et les process déterminants dans vos fonctions. Nos événements réels et nos séminaires virtuels vous donnent accès à des commentaires et recommandations neutres par des experts sur les problèmes et défis que vous rencontrez quotidiennement. Notre communauté en ligne "IT Knowledge Exchange" (Echange de connaissances IT) vous permet de partager des questionnements et informations de tous les jours avec vos pairs et des experts du secteur.