

ACCÉLÉRER VOTRE BUSINESS AVEC SQL SERVER 2012

> Par Loïc Duval

La donnée est au cœur de tous les Business. Elle est tout à la fois, le contenu informationnel de l'entreprise, son socle de fonctionnement, son savoir, son quotidien, son potentiel de croissance, ... Dès lors, elle fait du SGBD une pièce centrale de son activité. D'autant qu'aujourd'hui, celui-ci est devenu bien plus qu'une machine à stocker des transactions. ETL, Data Warehouse, BI, Big Data, en sont des constituantes fondamentales qui sont, ou devraient, être intégrées en standard. Dans un tel contexte, un SGBD ancien, incomplet, ou mal assemblé peut se révéler un frein au bon fonctionnement de l'entreprise. De même, un SGBD modernisé, exploitant toutes les facettes évoquées plus haut, s'avère un véritable accélérateur de Business !

Bien plus qu'un SGBD, SQL Server 2012 s'affirme comme une vraie solution complète couvrant tous vos besoins en matière de données, du transactionnel au Big Data, en passant par l'analytique, le décisionnel et l'ETL. Pour Damien Cudel, Chef de Marché Plateforme Applicative chez Microsoft France, ce large spectre fonctionnel proposé en standard constitue le point fort de la solution : « Si on veut aider les clients à accélérer leur Business, il faut être en mesure de leur proposer des technologies tout à la fois complètes (couvrant toutes les couches du transactionnel à l'analytique, du serveur jusqu'au poste de travail), intégrées (compatibilité applicative et commodity hardware), familières (exploitant des outils largement déployés, déjà maîtrisés des utilisateurs car s'inscrivant dans Office, insérés dans le corpus des connaissances des IT), et bien évidemment rapides (à la pointe des technologies et clés en main) ».



Accélérer l'exécution

Tordons immédiatement le cou à une légende urbaine qui veut que SQL Server, qui s'appuie exclusivement sur les architectures x86 (32 et 64 bits), n'ait pas la carrure nécessaire pour supporter les Business les plus exigeants. Si l'on regarde la diversité des entreprises qui ont opté pour SQL Server pour soutenir leur Business et leur croissance, on constate que le SGBD de Microsoft gère en production chez un client des bases de plus de 100 To, chez un autre une table de plus de 1,5 trillion de lignes (1500 milliards de lignes), chez un autre une application SQL s'appuyant sur 92 Peta-octets de données. Aujourd'hui, avec un seul et unique serveur, on peut gérer 130 000 transactions par seconde. L'appliance HP UniDB réalise 650 000 E/S par seconde! A l'autre bout du spectre, certaines implémentations matérielles proposées par les constructeurs partenaires aspirent plus de 18 Go par seconde!

Mais, au-delà de ses performances pures, SQL Server propose aussi des technologies spécifiquement conçues pour booster les requêtes, la génération des cubes ou l'exploration des données. Autant d'accélération qui rendent l'entreprise plus agile car plus réactive et productive. Elles permettent non seulement de traiter des volumétries supérieures mais aussi de mettre en œuvre des analyses en temps réel de certaines données. Parmi elles, la réduction de l'impact des E/S via les technologies In Memory et une meilleure distribution des charges via la parallélisation massive des traitements constituent deux avancées majeures.

Le In Memory et SQL Server

Les technologies « In Memory » s'insinuent au fur et à mesure à tous les étages de SQL Server. Cubes OLAP et analyses multidimensionnelles tirent ainsi profit du moteur xVelocity Analytics Engine. Introduit à l'origine avec PowerPivot pour Excel 2010, ce moteur est désormais intégré au cœur de la BI de SQL Server. Il n'est nullement besoin d'acquérir des technologies complémentaires ou du matériel spécialisé. Mixant des techniques comme le stockage en colonne, les caches « en mémoire » et des algorithmes d'agrégation hautement parallélisés, il accélère considérablement les requêtes analytiques et permet aux utilisateurs d'interagir en temps réel avec les modèles et des millions de lignes de données. xVelocity Analytics est ainsi capable de scanner environ 20 milliards de lignes par seconde! Autre innovation, les ColumnStore Index accélèrent par un facteur de 10 à 100, la plupart des requêtes types

FAST TRACK & APPLIANCES SQL SERVER : SIMPLES, RAPIDES ET SANS RISQUES

Usages	Architecture de Références	Appliances
Data Warehousing	Fast Track (HP, Dell, IBM)	HP Business Data Warehouse (HP BDW < 5To) Dell QuickStart DW 1000 (<5To) & 2000 (<12To) HP Enterprise Data Warehouse (basé sur SQL PDW) Dell PDW (basé sur SQL PDW)
Business Intelligence	BULL Fast track	HP Business Decision Appliance (BDA) (Sharepoint + PowerPivot + SQL Server)
SQL Server for Private Cloud (DB Consolidation)	HP Database Consolidation RA	HP Virtualized System
Transactionnel (OLTP)	HP Enterprise OLTP (2.3 milliards tx/heure)	

réalisées sur un Data Warehouse. Ainsi, grâce à xVelocity et les ColumnStore Index, la banque de Nagoya a réduit à 2 secondes le temps d'une extraction de 100 millions d'enregistrements qui lui prenait préalablement 30 minutes !

Parallélisme des traitements

Si le « In Memory » est une approche efficace aux problématiques de performances, elle n'est pas la seule. La parallélisation des traitements joue aujourd'hui un rôle essentiel notamment dans l'exploitation des flux d'informations (en provenance du Web, des réseaux sociaux ou des échanges en bourse) ainsi que dans les traitements sur les entrepôts de données. Dans le premier cas, SQL Server intègre en standard un moteur CEP (Complex Event Processing) dénommé StreamInsight. Il est indispensable pour gérer en temps réel, au niveau même du SGBD, des flux d'événements toujours plus nombreux. Il simplifie la mise en œuvre de tableaux de bord et de systèmes d'alertes s'appuyant sur des informations constamment mises à jour auxquelles il faut réagir en temps réel sans impacter les tâches de production. Dans le second cas, SQL Server 2012 Parallel Data Warehouse (PDW v2) est une appliance pour Data Warehouse évolutive, très performante et à faible coût, basée sur une architecture MPP (Massive Parallel Processing). Une approche qui permet de gérer efficacement des volumétries de plusieurs péta-octets, en parallélisant de façon automatique la répartition des données et l'intégralité de la chaîne de traitements.



Accélérer le déploiement

Le déploiement de nouvelles solutions, de nouveaux outils, de nouveaux projets, ne doit pas non plus être un frein à la volonté d'accélérer son Business. L'une des forces de SQL Server est de permettre, aujourd'hui, à vos projets de naître en limitant les risques et de passer en production plus rapidement. A l'heure où bien des entreprises cherchent à gagner

en réactivité et en souplesse au travers du Cloud, SQL Server s'impose comme la base de données offrant la plus grande souplesse de migration entre « On Premises » et « Cloud ». Windows Azure SQL Database offre un très haut niveau de compatibilité avec SQL Server : même langage T-SQL, même philosophie, mêmes structures, mêmes outils d'administration. Des technologies de réplication et de synchronisation intégrées simplifient le déploiement des bases dans le Cloud et la concrétisation de scénarios hybrides où seules les données utiles sont exposées dans le nuage et les données confidentielles protégées en local. Toutefois, le Cloud est loin d'être aujourd'hui la réponse universelle aux problématiques de déploiement. Pour les accélérer également « On Premises », Microsoft propose à la fois des outils de consolidation, des références d'architecture types (appelées Fast Track) et des appliances.

Consolider

Le nombre de bases de données a explosé en même temps que le nombre d'applications métier. Le besoin de consolider les serveurs pour simplifier leur administration se retrouve aussi au niveau de SQL Server et de l'administration des bases. Mais l'approche n'est pas nécessairement identique. La plupart du temps, au fil des nouveaux projets, les entreprises ont accumulé les serveurs dédiés mono-instance. Au travers de la gestion « multi-instances » (on installe plusieurs instances au sein d'une même machine physique) et de l'apparition des Contained Database (qui simplifient la gestion des sécurités), SQL Server 2012 offre une alternative à la simple virtualisation des serveurs SGBD (qui reste bien évidemment possible). Cela revient à une mutualisation des ressources avec une approche administrative différente. Un outil comme Microsoft MAP peut, d'ailleurs, vous proposer différents scénarios de consolidation et accélérer les phases de migration vers la dernière version de SQL Server.

Fast Track et Appliances

L'objectif de ces appliances, le pourquoi de leur existence, est de réduire le Time To Market, de réduire les risques, et d'optimiser le TCO avec un investissement sur une configuration équilibrée et calibrée spécialement pour le projet ou l'usage visé. Les architectures de référence et les Appliances SQL Server présentent des solutions optimisées et calibrées pour chaque scénario : Data Warehouse, Business Intelligence, OLTP (Transactionnel), Consolidation. Elles permettent d'accélérer les déploiements et d'assurer des performances optimisées selon la volumétrie considérée. Fast Track est à la fois une référence d'architecture et un label pour du Datawarehouse SMP qui offre au client la garantie que le constructeur et Microsoft ont travaillé ensemble pour bâtir une plateforme pré-configurée et optimale. HP, Dell, Bull, Cisco et IBM proposent aujourd'hui de telles solutions FastTrack pour SQL Server 2012. Les Appliances PDW sont des solutions « clés en main » de Datawarehouse MPP, proposées par HP et Dell, et destinées à accueillir des entrepôts avec des requêtes très complexes non prédictives sur des volumétries de quelques To à plusieurs Po.



Accélérer l'appropriation

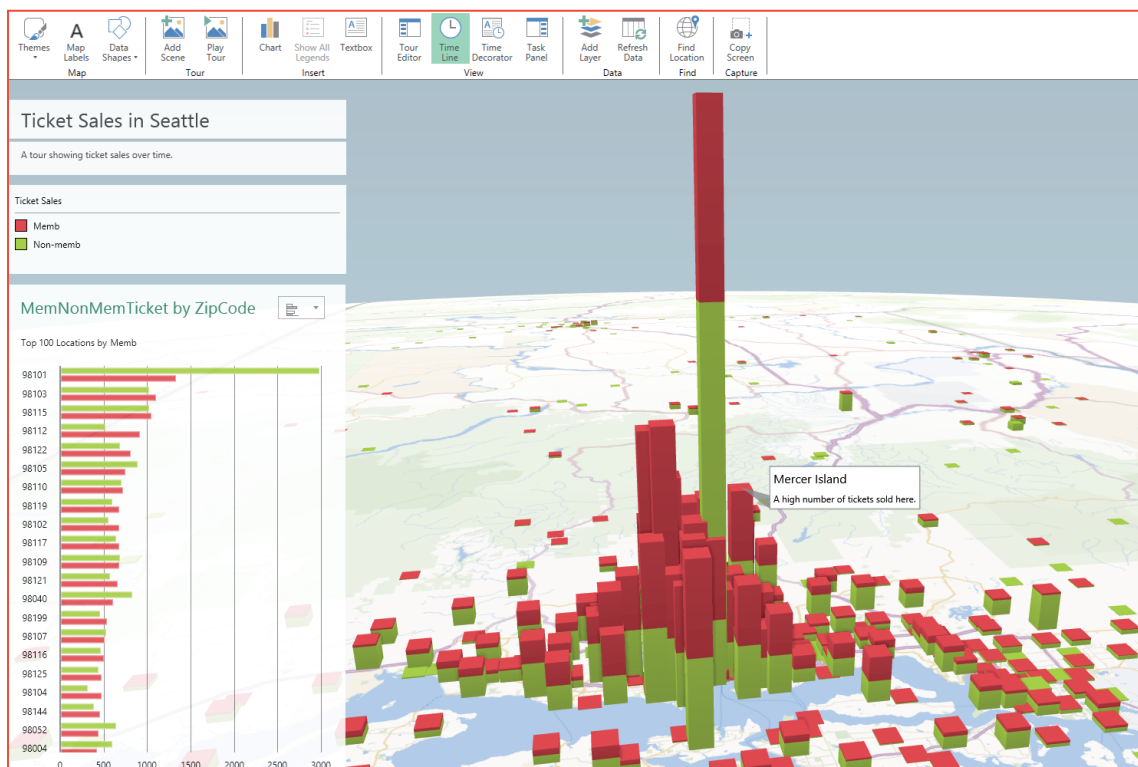
Il ne suffit pas de souhaiter accélérer son Business et de mettre en œuvre de nouveaux outils pour que cette volonté s'accomplisse. L'un des freins typiques à toute introduction de changements est la difficulté avec laquelle les collaborateurs peuvent se les approprier. Accélérer le Business, c'est aussi accélérer l'appropriation des outils que l'on met en place. Une appropriation qui doit être aussi simple, naturelle et rapide pour les IT que pour les utilisateurs.

De l'administrateur...

Côté IT, que l'on parle On Premises ou Cloud, que l'on parle OLTP ou BI, que l'on parle In Memory ou distribution multi-tiers, que l'on parle de réplication et auto-maintenance ou de PRA et haute-disponibilité, ce sont toujours les mêmes connaissances, les mêmes compétences et les mêmes outils qui sont utilisés et mis en œuvre, avec une extrême simplicité intrinsèquement liée à Windows. Cette volonté de permettre à tout un chacun d'exploiter ses connaissances et compétences actuelles pour maîtriser les technologies de demain s'affirme même dans la diversité de technologies qui coexistent autour du Big Data. Avec SQL Server 2012 PDW, Microsoft introduit un nouveau moteur, dénommé Polybase, qui permet aux développeurs et DBA d'interroger en T-SQL aussi bien des bases relationnelles SQL Server que des tables NoSQL, non relationnelles, hébergées sous Hadoop. Une intégration transparente qui permet d'enrichir les analyses en combinant des informations diverses, structurées comme non structurées, et de gérer tous types de données avec les compétences acquises, sans avoir à maîtriser MapReduce ou des outils radicalement différents d'un SGBD.

... à l'utilisateur

Côté utilisateurs, Microsoft est sans aucun doute l'un des acteurs du marché qui a le plus œuvré à la démocratisation de la BI et à la généralisation du concept de Self-BI. Un concept qui cherche à offrir à tous les collaborateurs de l'entreprise un moyen d'explorer le patrimoine informationnel pour en découvrir la richesse cachée au cœur des données et en déduire de



Visualisation géographique des données avec GeoFlow pour Excel 2013

nouveaux débouchés. Encore faut-il que cette self-BI puisse facilement être auto-appropriée, sans nécessité de connaissances préalables en BI ou en programmation. L'idée fondatrice de la BI Microsoft est donc de ramener l'utilisateur sur des référentiels connus. D'où la volonté de faire d'Excel le support de cette « self-BI ». Une approche très logique lorsque l'on sait que la plupart du temps, les utilisateurs de SGBD ou BI concurrents passent leur temps à exporter les informations pour les intégrer et les manipuler sous Excel. Avec la BI Microsoft, tout est intégré en standard et tout est intégré directement dans Excel ! Les utilisateurs travaillent directement dans leur environnement favori et n'ont pas à jongler entre les outils. Et cette BI personnelle est désormais vraiment accessible à tous dans l'entreprise au travers d'outils très visuels tels que Power View (idéal pour raconter des histoires autour des données et démontrer visuellement différents scénarios au travers d'animations), GeoFlow (pour représenter en 3D toutes données géo-localisées) ou même Data Explorer (une nouvelle extension Excel d'exploration et de découverte d'informations).



Accélérer l'innovation

C'est désormais une évidence, nous sommes entrés dans une ère nouvelle. Celle des flux d'informations permanents, fruit de l'explosion des devices mobiles et des capteurs embarqués et de l'explosion des données non structurées. Les chiffres parlent d'eux-mêmes : 12 To de tweets sont échangés chaque jour, eBay brasse 100 Po d'informations, Facebook stocke plus de 10 milliards de photos, près de 240 To de données sont générées par un Boeing 737 au cours d'un vol transatlantique !

Innover grâce aux données

Tout change, y compris notre façon de travailler. Un changement propice aux innovations. L'entreprise est contrainte de se poser de nouvelles questions sur les usages induits par la mobilité, le Cloud et les réseaux sociaux et sur les services à mettre en œuvre pour s'y

adapter. Parallèlement, la BI influence la façon dont nous découvrons et collaborons autour des données. « Pourtant, la BI est encore sous utilisée en entreprise » constate Damien Cudel. « Seulement 13% des entreprises exploitent intensivement des outils d'analyse prédictive. En favorisant le recoupement des multiples sources, elle permet souvent de réagir plus vite, d'envisager de nouvelles opportunités avant ses concurrents, de dénicher de nouvelles idées ». On entre là au cœur même du Big Data. Une problématique qui ne se limite pas à des problèmes de volumétrie, mais s'étend à l'usage que l'on en fait et un potentiel qu'on en retire.

Incontournable Big Data

SQL Server constitue un socle fondamental face à ces besoins d'innovation et face aux 3 défis du Big Data :

- * le défi Volumétrique : avec des capacités d'extension vers le cloud, la parallélisation massive des traitements de PDW, les accélérations In Memory, les compressions délivrées en un clic, mais aussi avec ses outils de Data Management pour créer des référentiels ainsi que nettoyer et corriger les données.

- * le défi Vitesse : avec la capacité à absorber de nouveaux projets grâce à l'élasticité du Cloud tout en minimisant les prises de risque, ou encore la capacité à réagir en temps réel à la multitude d'évènements parallèles via StreamInsight.

- * le défi Variété des données : avec son moteur d'indexation Full Text intégré, son moteur d'analyse textuelle et bien évidemment l'intégration transparente des tables non relationnelles au travers de Polybase. Aujourd'hui, SQL Server compte bien vulgariser le Big Data comme il a vulgarisé la BI auprès des entreprises de toutes tailles. HDInsight Server (et son équivalent Cloud sur Windows Azure) est une implémentation 100% compatible d'Apache Hadoop offrant une connectivité naturelle avec Office Excel 2013 et les différents outils de BI et de reporting de SQL Server. Associée à Polybase, elle permet de tirer la substance moelle des différents flux d'informations quelles que soient les connaissances des administrateurs et des développeurs. ■

HEKATON

Il n'y a pas que le décisionnel qui bénéficie des accélérations « In Memory ». Microsoft a déjà dévoilé l'arrivée du moteur « Hekaton » pour le transactionnel. Si certains concurrents proposent des moteurs spécialisés qui placent généralement l'intégralité des bases en mémoire, Microsoft adopte une approche plus pragmatique puisque Hekaton est intégré au cœur même de la prochaine version de SQL Server. « C'est dommage d'avoir à tout mettre en mémoire. Il y a typiquement des tables qui n'ont pas besoin d'être placées en mémoire, car elles sont très peu exploitées. Or, la plupart des solutions In Memory du marché montent toutes les bases et toutes les tables en mémoire. Quelque part, on paye donc de la mémoire pour rien, finalement » explique ainsi Damien Cudel de Microsoft France. « L'approche de SQL Server avec Hekaton est différente et se montre à la fois plus logique, plus évoluée, et plus économique. Rien qu'en plaçant quelques tables clés en mémoire, on peut atteindre 90% des performances que l'on aurait eu en positionnant tout en mémoire sans gonfler son budget et sans changer les serveurs. ». Pour déterminer quelles tables auraient tout intérêt à être placées en mémoire - pour une accélération instantanée et totalement transparente des requêtes et de votre Business - Hekaton sera livré avec un outil repérant les points chauds et suggérant les mises en mémoire les plus impactantes dans le gain de performances.