

Teradata marie datawarehouse, in-memory et Big Data

Nous le présentions tandis que se déroulait Teradata Universe 2013 (lire « [Teradata Universe 2013 : l'opportunité Big Data](#) »), le géant des entrepôts de données rongait son frein avant d'annoncer l'adoption du in-memory.

Les troupes se sont rabattues sur le Big Data et la nouvelle fabric Data Warehouse 6700 (lire « [Teradata Universe 2013 : Fabric Based Computing et Data Warehouse 6700](#) »), mais il flottait dans l'air un parfum de base de données traitées en mémoire. Logiquement, les architectures d'analyses en temps réel s'accommodent particulièrement bien avec le traitement in-memory.

Teradata Intelligent Memory

Il n'aura donc fallu que quelques semaines à Teradata pour annoncer sa première architecture in-memory. Basée sur l'infrastructure Teradata Unified Data Architecture (lire « [Teradata Universe 2013 : Unified Data Architecture](#) »), la nouvelle architecture de Teradata qui associe l'intégration des données du datawarehouse Teradata à l'analytique Aster via une plateforme Hadoop, **Teradata Intelligent Memory** réunit la technologie de datawarehouse, le in-memory, le Big Data, et l'analytique Aster sur Hadoop.

Destinée à traiter de très gros volumes de données – la solution n'est pas signée Teradata pour rien ! -, la technologie in-memory permet de placer en mémoire les données les plus chaudes, donc les plus utilisées (selon Teradata, 90 % des requêtes ne concernent que 20 % des données), afin d'en assurer le traitement, la mise à jour et la synchronisation. Et cela sans passer par le goulot d'étranglement des I/O (entrées/sorties) puisque la mémoire est directement en contact avec le CPU (processeurs). Cette proximité, doublée d'une approche d'identification prédictive de la donnée chaude, évite les appels aux applications, aux requêtes SQL et aux ressources de stockage.

La température des données

L'approche globale de l'entrepôt de données par Teradata permet d'associer différents niveaux de température à la donnée. Les données chaudes, c'est à dire celles qui sont les plus utilisées, sont placées au plus proche du compute, c'est à dire généralement sur des disques durs rapides. Les données les plus froides, donc les moins utilisées, sont placées sur des disques durs moins coûteux, mais également moins rapidement accessibles.

Cette hiérarchisation de la données est bouleversée par les apports des technologies les plus récentes, qui permettent d'adopter aujourd'hui une stratégie de 'multi-températures'. Une première étape a été franchie avec l'intégration des disques SSD (*solid-state drive*) sur lesquels sont stockées les données chaudes. Les données non-structurées et les données annexes (qui généralement ne sont pas prises en comptes par les outils analytiques dédiés aux données structurées accessibles via des requêtes SQL) sont désormais exploitées et prennent place dans

l'échelle des température via les technologies de Big Data Hadoop. Et c'est maintenant au in-memory de s'inviter pour élever le premier niveau de température de la donnée.

Teradata propose également une technologie de migration de la donnée d'un niveau de stockage à l'autre et paramétrable avec **Teradata Virtual Storage**.

Datawarehouse, Big Data et in-memory

Le mariage du datawarehouse, du Big Data analytique et du in-memory s'impose comme une évidence. La partie entrepôt de données assure le support de très gros volumes de données gérés par des algorithmes sophistiqués pour dater, suivre et classer automatiquement la donnée stratégique.

Le Big Data Hadoop est la plateforme de consolidation et d'analyse en temps réel des données moins/non stratégiques stockées sur des ressources hétérogènes. Enfin, le traitement en mémoire des applications et des données chaudes est un accélérateur des performances analytiques, de la requête à la prise de décision.

Voir aussi

[Silicon.fr étend son site dédié à l'emploi IT](#)

[Silicon.fr en direct sur les smartphones et tablettes](#)