

Amazon DynamoDB, la base de données NoSQL en mode cloud... à la puissance SSD

Avec DynamoDB, son nouveau service managé de base de données NoSQL, Amazon permet à ses clients *cloud*, en quelques clics, de créer, dimensionner et piloter une table de base de données NoSQL, sans dégradation des performances, et avec un accès à la mesure de l'utilisation et des performances du service. Les données sont stockées sur disques SSD (*Solid State Drive* ou disque en mémoire flash), ce qui permet de réduire la latence des lectures/écritures à quelques millisecondes. La base bénéficie d'une réplique synchrone sur les *datacenters* régionaux AWS (*Amazon Web Services*) Availability Zones.

NoSQL n'est pas « non SQL »

Mais commençons par la base, le NoSQL (*Not only SQL*). Derrière ce nom se cache la difficulté qu'éprouvent certaines entreprises à disposer de bases de données massivement non structurées performantes, en particulier celles par lesquelles transitent ou qui stockent d'énormes quantités de données. Ici, le traitement de la donnée ne nécessite pas l'usage de l'ensemble des fonctionnalités des bases relationnelles. Une base de données NoSQL est donc plus simple qu'une base SQL, même si elle en reprend les fondamentaux, ce qui la rend plus performante dans les environnements gourmands en ressources, en particulier dans les services internet.

La performance matérielle indispensable

AWS vient rappeler que même dans le *cloud* l'équipement déployé par les fournisseurs de service demeure essentiel. En effet, Amazon DynamoDB ne pourrait exister si le service ne s'exécutait sur un système de stockage SSD. Les volumes, la diversité et l'intensité des échanges (I/O) des données stockées dans une base NoSQL imposent de disposer d'une infrastructure largement plus dimensionnée qu'une architecture de base de données classique. Amazon l'a bien compris et DynamoDB se présente comme la première architecture de stockage SSD proposée en tant que service.

Une couche d'analytique EMR

AWS n'a pas oublié non plus qu'une base de données uniquement centrée sur le stockage n'offre qu'un service limité. La capacité de trier, d'extraire et d'analyser les contenus est aujourd'hui essentielle. Pour Amazon, cela se traduit par l'intégration d'Amazon EMR (*Elastic MapReduce*), le *framework* Hadoop proposé sur le modèle « *pay-as-you-go* », une solution qui permet à la fois d'accéder aux données stockées en de multiples lieux (Amazon S3, RDS ou DynamoDB), mais surtout de disposer d'une solution d'analyse des données qui en stockant les résultats sur Amazon S3 réduit sensiblement le risque de corrompre la base DynamoDB.

Dernier détail – qui n'en est pas un ! – Amazon DynamoDB est proposé au prix de 1 dollar le gigaoctet par mois, ce qui offre le stockage à la puissance du SSD à un prix particulièrement attractif.

