

IT Press Tour : NuoDB, le futur de la base de données

Boston, Massachusetts – NuoDB... Que de chemin parcouru depuis notre rencontre avec ce projet qui s'appelait alors NimbusDB. À commencer par son changement de nom, qui pour faire face à la vulgarisation de l'expression « Nimbus » s'appelle désormais **NuoDB**.

« *Ce fut une importante décision, qui correspond également à l'importance du cloud* », commente **Barry Morris**, fondateur et CEO de NuoDB. Changement également dans les locaux qui hébergent la start-up, devenus spacieux. Lever 12 millions de dollars, ça aide ! Changement enfin dans le projet, qui désormais, en bêta 8, approche de sa finalisation.



Aux origines de NuoDB

Mais revenons tout d'abord sur les origines du projet. En commençant par le constat d'un marché riche, mais qui évolue peu technologiquement. La consolidation du marché des bases de données, naturelle en soit, a eu lieu il y a 5 à 10 ans.

Il reste très lié au datacenter avec un hébergement dédié et une capacité qui certes s'expriment aujourd'hui en téraoctets et en dizaines de milliers d'utilisateurs sur le modèle du client-cluster, mais qui semblent atteindre leurs limites.

Autre source du projet, la base de données NuoDB est basée sur les travaux de **Jim Starkey** – un des piliers de DEC, puis fondateur de Netfrasture, qui deviendra le cœur du moteur de stockage de MySQL – sur l'approche scientifique ACID (atomicity, consistency, isolation, durability) associée à la fiabilité des transactions dans la base de données.

Cela se traduit par un modèle de base de données web et dans le cloud, que nous décrivons [en seconde page](#).

Les temps changent !

Voilà qu'arrivent de nouveaux projets de bases de données. Mais que se passe-t-il sur la donnée ? Barry Morris, constate trois changements majeurs :

- **L'émergence de nouveaux datacenters**, avec de nouveaux modèles de stockage en

mode scale-out et une multiplication des machines qui entrent dans la périphérie du datacenter. « *La couche de présentation est devenue folle, comme la couche stockage. La base de données traditionnelle fonctionne sur une machine. Pour la faire évoluer il faut passer sur une machine plus puissante et générer des ruptures. C'est suffisant pour changer les règles. C'est un problème technologique !* »

- **Le big data**, qui répond à la taille des données dans le datacenter, « *qui est devenue gigantesque... La donnée doit venir de sources automatisées, et non plus manuelles. Elle demande de nouveaux vecteurs.* »
- **Les workloads**, enfin. « *Les workloads sont différents, ils deviennent un problème et demandent à réduire la latence.* »

« *C'est la raison de la création des nouvelles bases de données. Et pourquoi nous allons vers la création de nouvelles grandes compagnies sur les bases de données.* »

Le patron de NuoDB ne cache pas ses ambitions et entend obtenir sa part d'un marché qui se chiffre en dizaines de milliards de dollars.

« *L'avenir n'est pas à SQL, au transactionnel, dont les technologies ne bougent pas, et dont les architectures ne peuvent grossir. Les start-ups apportent l'agilité au moment où tout devient plus 'data intensive'. Nous avons le potentiel de déplacer le goulot d'étranglement. Et nous allons évoluer vers une fragmentation de l'offre et des technologies. La question est : "jusqu'où se spécialiser ?"* »

Le modèle webscale

Pour NuoDB, la base de données de demain s'appuie sur trois piliers :

- la puissance de l'accès relationnel SQL ;
- la scalabilité du cloud pour son élasticité, la virtualisation, la disponibilité, sa géodistribution et le zéro administration ;
- la fiabilité ACID pour l'atomicité, la consistance, l'isolation et la durabilité.

Concrètement, cela se traduit par une base de données en pétaoctets à la demande, dans le cloud, avec une disponibilité totale sur du multisite actif-actif ; pouvant accepter des centaines de milliers d'utilisateurs ; 20.000 TPS et des workloads d'une grande complexité, hybrides, transactionnels, analytiques ; et supportant un développement agile.

Lorsqu'on l'interroge sur la solution généralement adoptée aujourd'hui consistant à partitionner les bases, la réponse de Barry Morris est directe : « *Le partitionning est une très bonne idée... si vous n'avez rien d'autre à faire que partitionner !* »

NuoDB préfère adopter une autre approche, la création d'une **architecture émergente**, dans le sens de réunir des entités simples pour former un environnement complexe et partagé.

Une approche qui permet de multiplier les performances lorsque l'on ajoute des ressources. « *'Emergent' est la caractéristique d'être intelligents qui se réunissent, mais où personne ne commande. C'est la seule voie pour que la nature progresse. Lorsque l'on imagine une architecture émergente pour le système, centraliser l'information, le contrôle, le 'master directory' n'est pas la solution. Tout doit être en peer to peer, sans master data, sans directory central.* »

Le modèle technologique NuoDB

La base de données NuoDB repose sur un gestionnaire de stockage, **NuoDB Storage Manager**, qui donne accès aux données stockées dans le cloud, en local, ou en mode hybride ; un moteur de transactions, **NuoDB Transaction Engine**, le nœud sur lequel s'exécutent les transactions ACID ; et un broker **NuoDB Broker** qui assure la relation avec les applications et le load balancing.

L'ensemble est qualifié d'ATOM. Il répond à l'approche de l'atomicité du 'A' de ACID en apportant la sécurité de la « transaction atomique », dans laquelle 'toutes' les opérations sur la base de données s'exécutent ou ne s'exécutent pas.

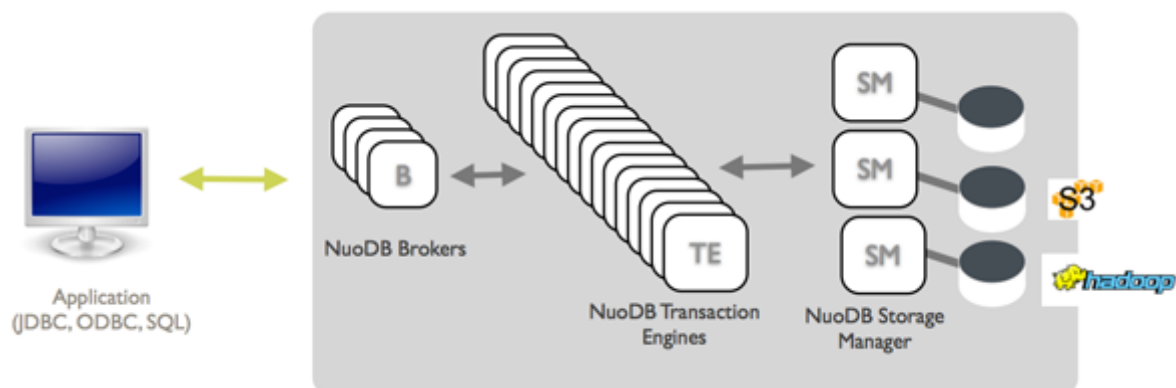
Dans un environnement NuoDB, les transactions (moteur de transactions) s'exécutent dans le Tiers 1 In-memory, tandis que le stockage des données est placé dans un Tiers 2, qui peut par exemple être Amazon ou Big Data/Hadoop. Ces deux Tiers ne doivent pas être sur une même machine, ce qui offre la capacité de disposer de plusieurs localisations des Tiers.

L'architecture de sécurité est complexe, le système de monitoring très développé, ce qui permet la très forte automatisation sans laquelle la solution ne pourrait s'exécuter.

Pour faire évoluer la base, il suffit d'ajouter des nœuds NuoDB Broker, Transaction Engine et Storage Manager. La scalabilité est totale jusque sur les I/O (entrées/sorties).

NuoDB affirme que la disponibilité est de 100 %, puisque les données, où qu'elles soient stockées, sont géodistribuées en actif/actif. Il n'y aurait donc pas de point de rupture.

Et le coût d'administration est réduit par l'usage d'une console unique de gestion. Enfin, l'accès à la donnée devient multi-tenant, les bases de données étant partagées dynamiquement entre les machines.



Un projet en phase bêta

NuoDB est actuellement en **version bêta 8**, avec 80 testeurs que l'éditeur entend porter à 100 avant de rendre son produit public.

Le mode de distribution fera certainement appel au modèle 'freemium' (l'accès au produit est gratuit dans une version complète, mais limitée en capacité) afin de la rendre largement accessible. Une première annonce sur ce que Barry Morris qualifie de « *produit révolutionnaire* » pourrait intervenir avant la fin de l'année.