

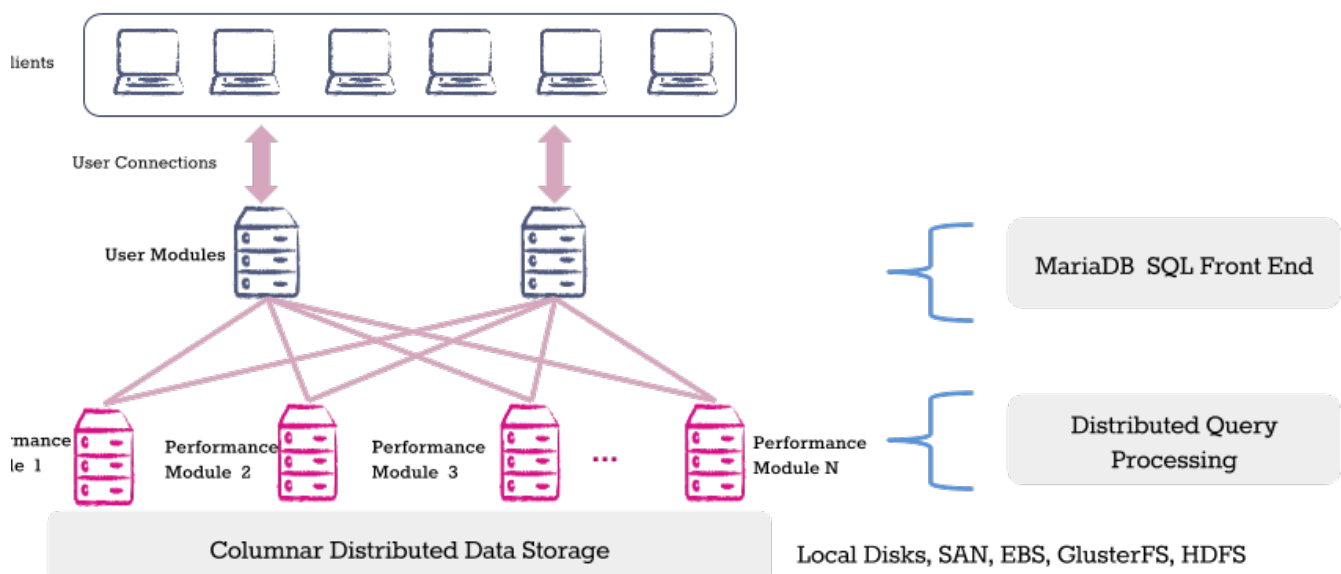
ColumnStore : MariaDB vient titiller le Big Data

MariaDB, le fork communautaire de MySQL, met sur le marché ColumnStore 1.0, un stockage de données en colonnes permettant aux utilisateurs de faire tourner en parallèle des processus analytiques (OLAP) et transactionnels (OLTP) depuis un même serveur MariaDB (en version 10.1). MariaDB cible clairement le remplacement des datawarehouses. La société promet ainsi un coût de revient annuel « *en moyenne 90% inférieur aux principaux systèmes d'entrepôt de données* ». Un comparatif qui repose sur les prix (élevés) des appliances dédiées, comme celles de Teradata ou Vertica. La solution ColumnStore peut être déployée dans le Cloud (elle est optimisée par AWS) ou sur un cluster local Linux.

MariaDB n'est évidemment pas la première base de données à proposer des applications analytiques sur des données de production, en quasi temps réel. Oracle, Microsoft, SAP, IBM ou encore Teradata offrent tous des solutions de ce type, ainsi qu'un certain nombre de jeunes sociétés comme VoltDB, NuoDB ou MemSQL. L'option est disponible dans les environnements MySQL depuis 2010 via une offre de l'intégrateur Percona.

Au-delà d'un million de lignes

ColumnStore, disponible en téléchargement, se décompose en deux services majeurs. Des User Modules d'abord, soit le front-end assurant l'orchestration des requêtes et auxquels sont connectés les utilisateurs. Et les Performance Modules, qui distribuent la prise en charge des requêtes. « *Et ces deux services sont scalables horizontalement* » afin d'assurer des performances élevées et une montée en charge du nombre d'utilisateurs, précise David Thomson, le vice-président en charge de l'ingénierie chez MariaDB dans un [billet de blog](#). « *Très peu de travaux de configuration et de maintenance sont nécessaires pour maintenir une performance élevée du système. Ainsi, il n'est pas nécessaire ni de constituer, ni de maintenir des index* », reprend le dirigeant.



Selon lui, si les moteurs OLTP classiques – InnoDB ou MyISAM – sont en mesure de traiter « *avec des performances raisonnables* » des requêtes portant sur moins d'un million de lignes, une fois ce cap franchi, une performance correcte sera difficile à obtenir et à maintenir dans le temps. D'où l'intérêt d'une alternative en colonnes, qui accélère l'accès aux données. Un point clef quand les jeux de données atteignent des millions ou des milliards de lignes.

ColumnStore en GPLv2

Pour faciliter l'accès à cette technologie, MariaDB propose un programme baptisé ColumnStore JumpStart qui fournit un environnement prêt à l'emploi sous 3 à 5 jours. ColumnStore, en réalité un fork d'InfiniDB (une entreprise qui a fait banqueroute en 2014), est proposé sous licence GPLv2. Une nouvelle qui devrait rasséréner (en tout cas un moment) la communauté des utilisateurs de la base de données Open Source : l'été dernier, MariaDB s'était attiré une volée de bois vert en passant la version 2.0 de sa technologie de proxy, MaxScale, sous licence propriétaire, la Business Source License (BSL) inventée par Monty Widenius, le créateur de MySQL aujourd'hui directeur technique de MariaDB.

BSL prévoit un basculement en Open Source à date fixe, avant laquelle l'utilisation de la technologie est soumise à des frais de licence. Si, dans les colonnes de *The New Stack*, David Thomson [explique](#) que MariaDB a décidé de placer les moteurs de stockage comme ColumnStore en Open Source, la société pourrait utiliser la BSL pour des technologies « *créant de la valeur au-dessus de ces moteurs* ».

A lire aussi :

[Pourquoi BlaBlaCar migre ses bases de données sous MariaDB](#)

[MariaDB : une version de printemps en toute sécurité](#)

Crédit photo : Jirsak / Shutterstock