

Résultats du test de performance de DataWalk

Questions

Résumé

Un problème courant avec tout type de base de données (par exemple, SQL, NoSQL) est que certaines requêtes complexes ne renvoient jamais de résultat. C'est pourquoi DataWalk a effectué des tests de référence pour mesurer les résultats des performances avec des requêtes profondes (c'est-à-dire de nombreuses jointures SQL), pour mesurer où une base de données SQL traditionnelle peut échouer et pour mesurer les performances de DataWalk avec un grand nombre de jointures.

Notez que si DataWalk est un système complet d'analyse de données et non une base de données, il comprend une puissante fonction permettant de générer facilement des requêtes en parcourant des ensembles de données et des connexions par le biais d'une interface visuelle.

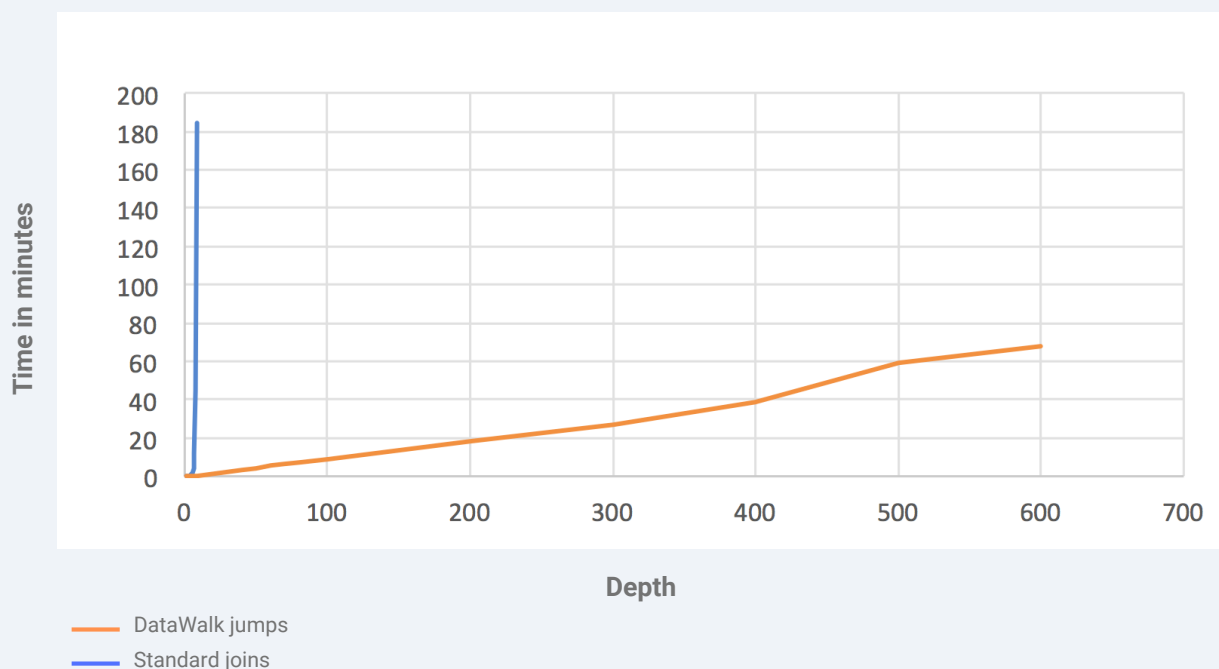
Ces requêtes peuvent ensuite être facilement

sauvegardées et modifiées, puis relancées instantanément en appuyant sur un "bouton facile" de DataWalk, ou elles peuvent être lancées à partir d'autres systèmes via l'API de DataWalk.

Les résultats de l'analyse comparative ont indiqué que pour un ensemble de données clients d'un échantillon, la base de données SQL n'était pas en mesure de renvoyer un résultat si neuf jointures étaient nécessaires. En revanche, DataWalk a fourni une performance linéaire à mesure que le nombre de jointures augmentait, jusqu'au test maximum de 600 jointures.

Cela montre qu'avec la technologie unique de DataWalk, non seulement les requêtes complexes seront rapidement calculées, mais toute requête donnera toujours un résultat, ce qui constitue un progrès significatif par rapport aux autres systèmes.

Depth traversing comparison



Le test

Ce test consistait à interroger les systèmes tout en augmentant la profondeur de l'interrogation en fonction des énoncés "inner join" et "left join". Les spécifications des requêtes ont été générées sur la base d'échantillons de rapports utilisés dans les environnements des clients.

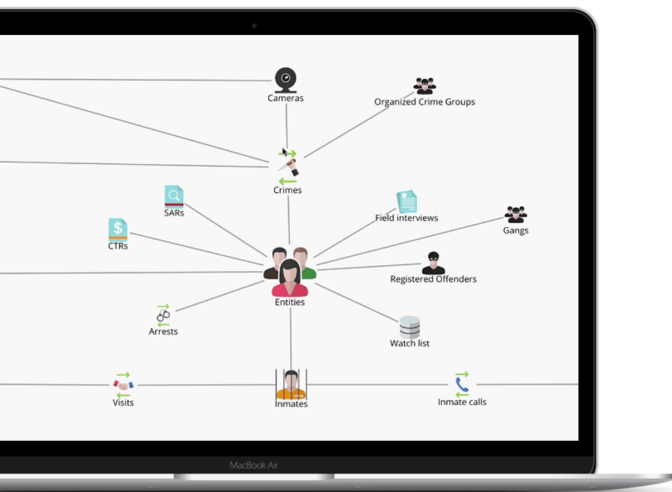
La configuration du système testé était identique pour les deux tests, avec huit serveurs de produits et une base de données de 1,3 To.

Résultats des tests

Le système basé sur SQL a montré une augmentation exponentielle du temps de réponse et de la consommation de ressources à mesure que le nombre d'entrées dans la base de données augmentait.

Pour cette charge de travail, le système SQL n'a pas pu effectuer plus de huit connexions. En essayant de joindre neuf structures, le système a déclaré des ressources insuffisantes après trois heures de traitement, et la tentative d'effectuer d'autres requêtes SQL n'a donné aucun résultat.

Sur le système DataWalk, les requêtes sont générées via le Universe Viewer, et la technologie DataWalk divise de manière transparente les requêtes complexes en plusieurs morceaux plus petits. Pour le même test mentionné ci-dessus, le logiciel DataWalk a été capable de joindre jusqu'à 600 structures avec une performance linéaire à mesure que le nombre de joints augmentait. Cela représente une amélioration spectaculaire de >>60X dans la capacité à effectuer des requêtes profondes et complexes.



DataWalk est une plateforme d'analyse de données complète permettant de révéler des modèles, des relations et des anomalies pour des données à grande échelle et provenant de plusieurs sources. DataWalk permet aux organisations d'importer et de fusionner rapidement de grandes quantités de données provenant de sources multiples dans une seule vue de données, puis d'analyser les données et les connexions via des interfaces visuelles simples.

Pour plus d'informations, voir www.datawalk.com.