

ÉTUDE DE CAS



albelli

Neo4j aide à monétiser les relations au sein de 1.2 pétaoctet de données d'un service de photos

SECTEUR

Logiciels / Imprimerie

CAS D'USAGE

Réseau social / Recherche basée sur les graphes

OBJECTIF

- Créer une nouvelle plateforme de stockage et de tri des photos pour remplacer la plateforme historique
- Mettre à disposition une solution de stockage bien organisée, sécurisée et simple à utiliser pour les images de tous les utilisateurs et faciliter le processus de sélection et de création de produits dérivés des photos

DÉFI

- Microsoft SQL Server impliquait des liaisons fastidieuses qui n'étaient plus adaptées aux volumes de données
- Développement complet en environnement de développement continu et mise sur le marché rapide

SOLUTION

1.2 pétaoctet de photos en local organisé et transféré dans le Cloud où Neo4j rassemble toutes les informations en rapport avec les méta données

RÉSULTATS

- Absorption d'une croissance continue de la clientèle sans souci de performance
- Gain de temps pour se consacrer aux objectifs stratégiques visant à une meilleure monétisation, la collecte des avis des clients et la réduction des coûts

L'entreprise albelli propose un logiciel qui synchronise et organise un volume de photos colossal pour plus d'un million d'utilisateurs. Grâce à Neo4j, son équipe met à disposition un outil moderne qui évolue en continu avec sa base de clients, génère davantage de revenus et lui permet de viser ses objectifs stratégiques de monétisation.

L'entreprise

Fondée en 2003, albelli combine les opportunités offertes par la nouvelle ère numérique en ligne et l'artisanat traditionnel de la reliure et de l'imprimerie. Sa solution en ligne de produits dérivés de photos a été primée à plusieurs reprises en Europe et s'adresse à différents marchés du continent. Leader du marché dans l'impression de livres photos, albelli fait partie des entre-prises de e-commerce qui se développent le plus vite en Europe et détient d'autres marques de produits photos telles que resnap, bonusprint, önskefoto et fotoknuden.

Le défi

albelli donne à ses utilisateurs la possibilité de synchroniser en un seul endroit des photos issues de différentes sources. Son logiciel organise automatiquement les photos en fonction de divers critères - date, heure, lieu par exemple - en événements affichés sur une frise chronologique simple à utiliser. Les clients évitent ainsi de devoir trier des milliers de photos. Ils gardent le contrôle total des options de partage et choisissent quels albums sont publics ou privés.

Une fois les photos téléchargées et organisées, les utilisateurs peuvent acheter différents produits tels des albums photos, calendriers, mugs et portraits sur toile.

Le nombre de photos est considérable - en moyenne 2000 photos par utilisateur, pour plus d'un million d'entre eux - tout comme la taille des fichiers combinés à tous les critères de description de chaque photo. Ainsi, l'entreprise devait gérer d'énormes volumes de données.

Elle utilisait encore Microsoft SQL Server, une base de données relationnelle lente et fastidieuse à utiliser à cause du grand nombre de liaisons qu'elle requiert pour les requêtes portant sur les relations. La nécessité d'ajouter à l'ensemble les capacités d'une autre base de données s'imposait pour surmonter ces défis de relations entre données.

L'équipe de développement est partie de zéro, sur le principe que leur domaine était constitué de relations semblables à des graphes entre photos et utilisateurs et sachant que traduire ce domaine similaire à des graphes sur un serveur SQL allait poser problème. Ils ont alors pris la décision de lancer un POC* avec Neo4j.

«Si le collectionneur typique de photos détient au moins 2000 photos, comment fait-on pour gérer des dizaines de nœuds sur chacune de ces photos tout en préservant la rapidité et la fluidité du produit ?» s'interroge Josh Marcus, Directeur Technique d'albelli. «C'est le plus grand défi technique que nous avons rencontré».

*Proof of concept

ÉTUDE DE CAS



«Le grand avantage avec Neo4j a été de pouvoir nous concentrer sur la modélisation de nos données et la façon de proposer le meilleur service à nos clients, plutôt que d'agoniser à chercher comment structurer les tableaux et les liaisons. De plus, il n'a fallu coder que très peu, ce qui nous a permis de nous consacrer à nos clients.»

– Josh Marcus
Directeur Technique,
albelli

La stratégie

L'entreprise pouvait compter sur une équipe de recherche et de développement agile, comptant seulement cinq personnes. Quand elle a commencé à élaborer la nouvelle architecture de données, elle a examiné différentes bases de données. Elle ne cherchait pas simplement une base pour stocker les données mais aussi le moyen d'exploiter les relations existantes entre les données.

Cette recherche l'a conduite directement à Neo4j. L'équipe a découvert cette solution lors d'une conférence et a pris contact pour en savoir plus. Une fois la décision prise de travailler avec Neo4j, l'équipe a tiré parti de tous les outils de formation mis à disposition par Neo4j - dont les formations en ligne et la 'coding academy' - pour finalement décrocher la certification professionnelle de Neo4j.

Grâce à une solide maîtrise de Cypher et de Neo4j, l'étape suivante a consisté à migrer leur important volume de données dans la nouvelle base de données de graphes.

La Solution

Au plus fort du projet, l'équipe a utilisé jusqu'à 700 instances EC2 exécutées à pleine vitesse pendant un mois et demi pour transférer vers un stockage Cloud 1.2 pétaoctet de données locales incluant 500 millions d'images et les données des clients. Cette migration a sollicité jusqu'à cinq serveurs Neo4j lourdement équipés et a concerné plus de 10 milliards de messages envoyés dans le système. Au cours de ce processus, albelli a également extrait les informations sur la date, l'heure et le lieu de chaque photo afin de pouvoir ensuite les relier entre elles.

«Le grand avantage avec Neo4j a été de pouvoir nous concentrer sur la modélisation de nos données et la façon de proposer le meilleur service à nos clients, plutôt que d'agoniser à chercher comment structurer les tableaux et les liaisons,» relate Josh Marcus. «De plus, il n'a fallu que très peu coder, ce qui nous a permis de nous consacrer à nos clients.»

Au final, leur base de données affichait un total d'un milliard de nœuds, 4.1 milliards de propriétés et 2.6 milliards de relations, Neo4j servant de base de données centrale. Outre Neo4j, albelli a utilisé d'autres technologies de bases de données dans son architecture, dont Redis, DynamoDB, Aurora et Microsoft SQL Server, ce qui lui permet d'exploiter une approche de persistance multi langage.

Les résultats

En un peu moins de deux mois, la petite équipe d'albelli a pu transférer ses données dans Neo4j et lancer une base de données plus rapide et réactive pour s'adapter à une base de clientèle en croissance. La solution lui fait gagner un temps considérable et lui permet de se consacrer à l'avenir, en particulier à de nouvelles modalités de monétisation, à la collecte des avis des clients et à la réduction des coûts.

Neo4j est le leader de la technologie des bases de données de graphes. Avec le plus grand nombre de déploiements au monde, Neo4j aide des entreprises mondiales - comme Airbus, [Michelin](#), [NASA](#), [Crédit Agricole](#) et [Volvo Cars](#) - à prédire et identifier la façon dont les personnes, les processus, les lieux et les systèmes sont interconnectés. Grâce à cette approche par les relations, les [applications](#) mises au point en utilisant Neo4j relèvent les défis associés aux données connectées, tels que [l'analytique et l'intelligence artificielle](#), la [détection de fraude](#), les [recommandations](#) en temps réel et les [graphes de connaissance](#). Pour en savoir plus, merci de consulter [Neo4j.com](#) et [@Neo4jFr](#).

Royaume-Uni uk@neo4j.com
France ventes@neo4j.com
Pays scandinaves nordics@neo4j.com
Allemagne vertrieb@neo4j.com
Europe du sud southern-europe@neo4j.com