

FALLSTUDIE



albelli

Neo4j ermöglicht Fotoservice die Monetarisierung der Beziehungen von 1,2 Petabyte Daten

BRANCHE

Software / Druck

ANWENDUNGSFALL

Soziale Netzwerke / Graphbasierte Suche

ZIEL

- Neue Plattform zur Speicherung und Organisation von Fotos, um die bestehende Plattform abzulösen
- Übersichtliche, benutzerfreundliche und sichere Speicherlösung für alle Bilder eines Benutzers für eine einfachere Auswahl und Erstellung von Fotoprodukten

HERAUSFORDERUNGEN

- Microsoft SQL Server erforderte umständliche JOINS, die angesichts des erreichten Datenvolumens nicht mehr tragbar waren
- Ganzheitliche Entwicklung in einer kontinuierlichen Deployment-Umgebung sowie kurze Time-to-Market

LÖSUNG

Übertragung und Organisation von 1,2 Petabyte Daten von lokaler Umgebung in die Cloud, wobei mit Metadaten verknüpfte Informationen in Neo4j enthalten sind

ERGEBNISSE

- Kontinuierliches Benutzerwachstum ohne Performanceverlust möglich
- Mehr Zeit für strategische Ziele wie Monetarisierung, Kundenfeedback und Kostensenkung

albelli unterstützt mit seiner Software über eine Million Benutzer bei der Synchronisation und Organisation großer Datenbestände und Fotos. Mit Neo4j hat das Unternehmen jetzt ein modernes Tool entwickelt, das gemeinsam mit seiner Benutzerbasis wächst, höhere Einnahmen erzielt und eine strategische Monetarisierung ermöglicht.

Das Unternehmen

albelli wurde 2003 gegründet und verbindet die Möglichkeiten des neuen digitalen Online-Zeitalters mit den „traditionellen Handwerkskünsten“ Buchbinderei und Druck. Die Software für Online-Fotoprodukte gewann mehrere Auszeichnungen in ganz Europa und bedient zahlreiche Märkte. albelli ist Marktführer im Fotobuchdruck und zählt zu den schnellst wachsenden E-Commerce-Unternehmen in Europa. Zu albelli gehören weitere Marken für Fotoprodukte wie resnap, Bonusprint, önskefoto und fotoknudsen.

Die Herausforderung

Mit albelli können Benutzer ihre Fotos aus mehreren Quellen an einem Ort synchronisieren. Die Software organisiert die Fotos automatisch anhand einer Vielzahl von Kriterien wie Datum, Uhrzeit und Ort und ordnet sie Ereignissen zu, die in einer anschaulichen Zeitleiste angezeigt werden. Kunden ersparen sich so das umständliche Sortieren und Organisieren Tausender Fotos. Dabei behalten sie jedoch die volle Kontrolle über die Freigabeoptionen und legen selbst fest, welche Alben öffentlich oder privat sind.

Nach dem Hochladen und Organisieren ihrer Fotos können die Benutzer eine Vielzahl von Produkten in Auftrag geben – von Fotoalben über Kalender und Tassen bis hin zu Portraits auf Leinwand.

Jeder der über eine Million Benutzer hat durchschnittlich 2.000 Fotos. Hinzu kommt die Größe der Dateien in Verbindung mit allen dazugehörigen Metadaten pro Foto. Dies zusammen erzeugt eine gewaltige Datenmenge.

Zur Verarbeitung dieser Daten nutzte die Software die relationale Datenbank Microsoft SQL Server. Mit wachsender Datenmenge erwies sich diese Datenbank wegen der vielen JOINS für beziehungs-basierte Abfragen als zu umständlich und zu langsam. Bald war klar, dass hier die Fähigkeiten einer anderen Datenbank gefragt waren, die die Herausforderungen in Sachen Datenbeziehungen meistern konnte.

Das Entwicklungsteam machte sich an die Arbeit und erkannte die graphenähnlichen Beziehungen zwischen Fotos und Benutzern. Gleichzeitig war man sich bewusst, dass eine Überführung dieser graphenähnlichen Domäne auf SQL Server kaum Aussicht auf Erfolg haben dürfte. Man entschied sich daher für ein Proof-of-Concept-Projekt mit Neo4j.

„Angesichts von durchschnittlich 2.000 Fotos pro Benutzer stellte sich uns die Frage, wie wir Dutzende von Knotenpunkten pro Foto handhaben konnten, ohne dass die Geschwindigkeit unseres Angebots darunter leidet“, erklärt Josh Marcus, Chief Technology Officer bei albelli. „Hier lag unsere größte technische Herausforderung.“

FALLSTUDIE



„Neo4j hat den großen Vorteil, dass wir uns auf die Modellierung unserer Daten und auf einen Top-Service für unsere Kunden konzentrieren können, anstatt uns mit der Strukturierung von Tabellen und JOINS beschäftigen zu müssen. Zudem war der Programmieraufwand gering, sodass wir unsere Kräfte auf unsere Kunden fokussieren konnten.“

– Josh Marcus
Chief Technology Officer,
albelli

Die Strategie

Das Forschungs- und Entwicklungsteam von albelli umfasst nur fünf Mitarbeiter. Als man mit dem Aufbau der neuen Datenarchitektur begann, standen zunächst verschiedene Datenbanken zur Auswahl. Doch es ging nicht allein um eine Datenbank zur Datenspeicherung, sondern auch um die Möglichkeit, die Beziehungen zwischen den Daten zu nutzen.

Diese Anforderung führte das Team direkt zu Neo4j. Aufmerksam auf die Software wurde man erstmals auf einer Konferenz. Nach der Entscheidung für Neo4j nutzte das Team alle verfügbaren Schulungswerkzeuge – einschließlich der Online-Schulungen und der Coding Academy – für die Weiterbildung zum Neo4j Certified Professional.

Ausgestattet mit guten Cypher-Kenntnissen und dem nötigen Wissen rund um Neo4j konnten die riesigen Datenmengen in die neue Graphdatenbank überführt werden.

Die Lösung

In der Spitze nutzte das Team bis zu 700 EC2-Instanzen, die eineinhalb Monate lang mit maximaler Geschwindigkeit liefen, um 1,2 Petabyte an Daten – darunter 500 Millionen Bilder plus Kundendaten – von den Systemen vor Ort in den Cloud-Speicher zu übertragen. Die Migration umfasste bis zu fünf leistungsstarke Neo4j-Server. Hierbei wurden mehr als 10 Milliarden Nachrichten übertragen. Gleichzeitig wurden Datums-, Zeit- und Ortsinformationen von jedem Foto extrahiert, um die Fotos untereinander in Beziehung setzen zu können.

„Neo4j hat den großen Vorteil, dass wir uns auf die Modellierung unserer Daten und auf einen Top-Service für unsere Kunden konzentrieren können, anstatt uns mit der Strukturierung von Tabellen und JOINS beschäftigen zu müssen“, so Josh Marcus. „Zudem war der Programmiereffort gering, sodass wir unsere Kräfte auf unsere Kunden fokussieren konnten.“

Letztendlich enthält die Datenbank insgesamt über eine Milliarde Knoten, 4,1 Milliarden Eigenschaften und 2,6 Milliarden Beziehungen, wobei Neo4j als zentrale Datenbank dient. Neben Neo4j setzt albelli auch andere Datenbanktechnologien in der Gesamtarchitektur ein, darunter Redis, DynamoDB, Aurora und Microsoft SQL Server. Das ermöglichte es, die Vorteile der polyglotten Persistenz voll auszuschöpfen.

Die Ergebnisse

In nicht einmal zwei Monaten konnte das kleine Entwicklungsteam von albelli seine Daten nach Neo4j übertragen und eine agilere und reaktionsschnellere Datenbank für einen wachsenden Kundenstamm einführen. Dabei gewann man wertvolle Zeit für zukunftsweisende strategische Aufgaben, wie beispielsweise neue Wege der Monetarisierung, des Kundenfeedbacks und der Kostensenkung.

Neo4j ist der führende Anbieter von Graphtechnologie. Die weltweit am häufigsten eingesetzte Graphdatenbank unterstützt Unternehmen wie [Deutsches Zentrum für Diabetesforschung e.V.](#), [NASA](#), [UBS](#) und [Daimler](#) darin, Zusammenhänge zwischen Menschen, Prozessen, Standorten und Systemen aufzudecken und datengestützte Vorhersagen zu treffen. Der Fokus auf Datenbeziehungen ermöglicht es, smarte Anwendungen zu entwickeln und die Herausforderungen vernetzter Daten zu meistern – von [Analytics und künstlicher Intelligenz](#) über [Betrugserkennung](#) und [Echtzeit-Empfehlungen](#) bis hin zu [Knowledge Graphen](#). Weitere Informationen unter [Neo4j.com](#).

Großbritannien uk@neo4j.com
Frankreich ventes@neo4j.com
Skandinavien nordics@neo4j.com
DACH vertrieb@neo4j.com
Südeuropa southern-europe@neo4j.com