

LIVRE BLANC

Top 10 des cas d'utilisation des bases de données de graphes

Révéler de nouvelles possibilités
avec les données connectées

Jim Webber, Chief Scientist, Neo4j

TABLE DES MATIÈRES

| | |
|--|----|
| Introduction | 1 |
| Cas d'utilisation n°1 : Détection de fraude | 2 |
| Cas d'utilisation n°2 : Moteurs de recommandation en temps réel | 4 |
| Cas d'utilisation n°3 : Graphes de connaissances | 6 |
| Cas d'utilisation n°4 : Lutte contre le blanchiment d'argent | 8 |
| Cas d'utilisation n°5 : Gestion des données de référence | 10 |
| Cas d'utilisation n°6 : Gestion de la chaîne d'approvisionnement | 12 |
| Cas d'utilisation n°7 : Gestion des opérations informatiques et de réseau | 14 |
| Cas d'utilisation n°8 : Traçabilité des données | 16 |
| Cas d'utilisation n°9 : Gestion des identités et des accès | 18 |
| Cas d'utilisation n°10 : Nomenclatures | 21 |
| Conclusion | 23 |

Top 10 des cas d'utilisation des bases de données de graphes

Révéler de nouvelles possibilités avec les données connectées

Jim Webber, Chief Scientist, Neo4j

Introduction

Chaque année, le volume « big data » ne cesse d'augmenter. Les chefs d'entreprise font face au défi de gérer des ensembles de données toujours plus volumineux tout en générant de l'information à partir de ces données. Comment les directeurs des systèmes d'information et de la technologie peuvent-ils produire cette information ?

Pour paraphraser [Seth Godin](#), les entreprises doivent cesser de se contenter de collecter des données et commencer à s'intéresser à la façon dont elles sont connectées. En d'autres termes, les [relations entre les points de données](#) deviennent presque plus importantes que les points de données eux-mêmes.

Les [systèmes de base de données relationnelle traditionnels](#) ne sont pas efficaces pour gérer les relations entre les points de données. En raison de leurs modèles de données tabulaires et de leurs schémas rigides, il est difficile d'y ajouter de nouvelles relations ou de nouveaux types de relations. Pour tirer parti des connexions entre les données, les entreprises ont donc besoin d'une technologie de base de données capable de stocker les informations à propos de ces relations en tant qu'entité de premier ordre. Cette technologie n'est autre qu'une [base de données de graphes](#).

[L'avenir réside dans la technologie des graphes](#). Les bases de données de graphes sont conçues pour un stockage efficace des relations entre les points de données et permettent également l'ajout de nouveaux types de relations ainsi que l'[adaptation du modèle de données](#) au fur et à mesure que les besoins métier évoluent.

Au-delà des seules bases de données, la technologie des graphes inclut également le domaine émergent de la [science des données de graphes](#) pour l'analytique prédictive et l'apprentissage automatique, ainsi que la [visualisation des données de graphes](#) pour l'identification et l'exploration de données.

Alors comment les entreprises peuvent-elles exploiter la technologie des graphes pour générer un avantage concurrentiel et une valeur métier à partir des données connectées ?

Découvrez le top 10 des cas d'utilisation des bases de données de graphes :

Cas d'utilisation n°1 : Détection de fraude

Chaque année, la fraude fait perdre des milliards d'euros aux banques et aux compagnies d'assurance.

Les méthodes traditionnelles de détection de fraude ne parviennent pas à limiter ces pertes car elles se base sur des analyses discrètes susceptibles de fausser les résultats. De ce fait, les fraudeurs opèrent de manière de plus en plus sophistiquée et mettent au point des méthodes variées pour exploiter les faiblesses de l'analyse discrète.

Les bases de données de graphes offrent de [nouvelles méthodes de détection des réseaux de fraude](#) et autres escroqueries complexes avec un niveau élevé de précision grâce à l'analyse contextuelle avancée des relations entre les données. Ainsi, les bases de données de graphes sont capables de stopper des scénarios de fraude élaborés en temps réel.

Pourquoi utiliser la technologie des graphes dans le cadre de la détection de fraude ?

Bien qu'aucune mesure de détection de fraude ne soit parfaite, il est possible d'introduire des améliorations considérables en prenant en compte les relations entre les points de données.

Comprendre ces relations et leur donner un sens ne signifie pas nécessairement collecter de nouvelles données. En effet, il est possible d'extraire des informations significatives à partir des données existantes en examinant le problème sous un nouvel angle : le [graphe](#).

ANALYSE DE LIENS COMPLEXES

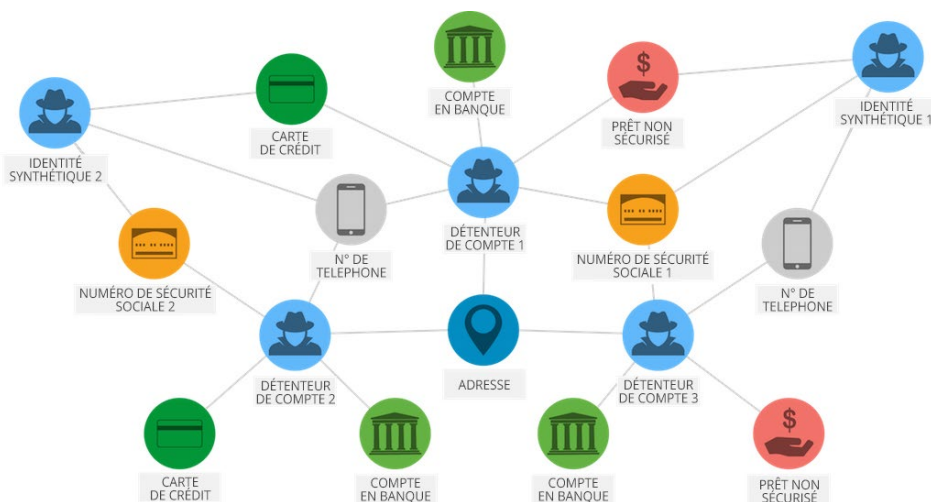
Révéler des réseaux de fraude nécessite d'effectuer une analyse des relations entre des données d'une grande complexité ; une difficulté exacerbée par la croissance des réseaux de fraude.

DÉTECTION ET PRÉVENTION DE FRAUDE

Pour lutter contre un réseau de fraude, il est nécessaire d'effectuer une analyse en temps réel des relations au sein d'un ensemble de données interconnectées, depuis la création du faux compte jusqu'à la transaction frauduleuse.

RÉSEAUX DE FRAUDE ÉVOLUTIFS ET DYNAMIQUES

Les réseaux de fraude ne cessent de se développer et les applications doivent être en mesure de détecter les schémas de fraude dans cet environnement en constante évolution.



Contrairement à la plupart des autres méthodes d'analytique, les graphes sont conçus pour exprimer les relations entre les données. Les bases de données de graphes révèlent des schémas qu'il est difficile de détecter avec les méthodes traditionnelles de représentation telles que les tableaux. De plus en plus d'entreprises ont recours aux bases de données de graphes pour résoudre une variété de problèmes liés aux données connectées, notamment la [détection de fraude](#).

Exemple : Allianz Benelux

Forte de plus de 2 000 employés, [Allianz Benelux](#) est une entreprise transfrontalière qui opère dans trois pays : la Belgique, les Pays-Bas et le Luxembourg. Allianz Benelux propose des solutions d'assurance dans tous les domaines, de l'assurance dommages à l'assurance vie et santé, sur un marché où prédominent les courtiers.

Allianz applique une tolérance zéro face à la fraude et jusque à présent, visualiser les comportements suspects avec la technologie des bases de données relationnelles était trop difficile.

Après une étude de marché approfondie, Allianz Benelux a choisi Neo4j comme plateforme de graphes en raison de sa nature évolutive, flexible, adaptée à l'entreprise et de sa prédominance sur le marché. Depuis, Allianz Benelux a rapidement adopté cette approche novatrice et puissante pour sécuriser l'avenir de ses clients, une réussite qui s'est directement traduit par des bénéfices commerciaux.

Avec la technologie des graphes, les activités potentiellement frauduleuses sont repérées dans l'écosystème d'Allianz grâce à une visualisation des relations illicites et dissimulées du fraudeur. En rassemblant toutes les données clients dans une base de données de graphes, Allianz Benelux identifie les véritables risques encourus et détecte les risques non couverts ou les chevauchements de polices d'assurance.

« En deux ans, un profit opérationnel d'une valeur de deux millions d'euros a été identifié. Cette valeur s'avère même être structurellement sous-estimée, » confirme Sudaman Thoppan Mohanchandralal, directeur des systèmes de données et d'analytique.

« Nous avons élargi notre vision client au-delà des polices d'assurance habituelles, y compris sur les courtiers, les bénéficiaires et les réclamations, » résume Dr. Jan Doumen, directeur de l'école d'expertise du bureau des données et responsable de la thématique stratégique 'Informations et perspectives clients et courtiers'. « Nous pouvons facilement identifier les clients qui détiennent des contrats avec différents courtiers et comprendre pourquoi ; identifier les clients qui passent leurs contrats d'un courtier à un autre et repérer les bénéficiaires de plusieurs polices d'assurance. Nous identifions les personnes qui apparaissent dans plusieurs réclamations, à première vue sans rapport, et au final, nous passons d'une analyse descriptive à une analyse vraiment prédictive. »

Conclusion

Dans le cadre de la détection de fraude basée sur les graphes, il est fondamental d'augmenter la capacité de détection avec l'analyse des relations. Ceci dit, deux points sont clairs :

- À mesure que les processus métier deviennent plus rapides et automatisés, les marges de temps pour détecter la fraude se réduisent, d'où le besoin accru d'une solution en temps réel.
- Les technologies traditionnelles ne sont pas conçues pour détecter les réseaux de fraude élaborés. Or, les bases de données de graphes sont capables de créer de la valeur en effectuant une analyse des relations entre les données.

La technologie des graphes est le moteur idéal pour mettre en place des solutions efficaces de détection de fraude. Qu'il s'agisse de réseaux de fraude, d'associations de malfaiteurs ou de criminels informés opérant seuls, la technologie des bases de données de graphes permet de révéler une grande variété de schémas de fraude en temps réel.

Cas d'utilisation n°2 : Moteurs de recommandation en temps réel



Qu'une entreprise opère dans le secteur de la vente au détail, des services, des médias ou des réseaux sociaux, offrir aux utilisateurs des recommandations personnalisées en temps réel est essentiel pour maximiser la valeur offerte aux clients et rester compétitif. Contrairement aux autres données métier, les recommandations doivent être de nature inductive et contextuelle afin d'être considérées comme pertinentes par l'utilisateur final.

Avec la technologie des graphes, il est possible d'associer le comportement de navigation d'un utilisateur avec ses données démographiques et son historique d'achat afin d'analyser instantanément ses choix actuels et lui fournir des [recommandations pertinentes](#) en temps réel.

Pourquoi utiliser la technologie des graphes pour les moteurs de recommandation en temps réel ?

La technologie clé pour permettre des recommandations en temps réel est une base de données de graphes, qui surpasse les autres technologies de base de données lorsqu'il s'agit de connecter des ensembles de données clients et produits.

La formulation de recommandations pertinentes en temps réel ne peut être effectuée qu'avec une base de données capable de comprendre les relations entre les entités, ainsi que la qualité et la stabilité de ces connexions. Seule une base de données de graphes permet le suivi efficace de ces relations en fonction des achats, des interactions et des avis des utilisateurs afin de fournir un aperçu des besoins clients et des tendances produits.

Les moteurs de recommandation basés sur les graphes adoptent souvent deux approches principales : l'identification des ressources / produits qui intéressent les individus ou l'identification des individus susceptibles de s'intéresser à une ressource / produit donné. Dans les deux cas, les bases de données de graphes établissent les connexions et les corrélations nécessaires pour apporter les résultats les plus pertinents pour l'individu ou la ressource en question.

GRANDES QUANTITÉS DE DONNÉES ET DE RELATIONS

Les algorithmes de recommandation les plus courants, comme le filtrage collaboratif ou le filtrage basé sur le contenu, s'appuient sur la traversée rapide d'ensembles de données hautement interconnectés et en augmentation constante.

RECOMMANDATIONS PERTINENTES EN TEMPS RÉEL

Apporter des informations contextuelles en temps réel relève du défi. C'est pourquoi la puissance d'un système de recommandation réside dans sa capacité à faire une recommandation en temps réel en incorporant l'historique immédiat.

NOUVELLES DONNÉES ET RELATIONS

Le niveau de précision et pertinence des recommandations s'améliorent à mesure que les points de données sont ajoutés. Cette croissance rapide force le système de recommandation à s'adapter à la fois aux contraintes actuelles et à venir.

Exemple : eBay

Leader du commerce en ligne, eBay travaille à [améliorer la façon dont les acheteurs recherchent des articles](#). Or, les moteurs de recherche et les moteurs de recommandation de produits actuels sont incapables de fournir ou de déduire les informations contextuelles pour une requête d'achat donnée, comme l'explique RJ Pittman, vice-président senior en charge des produits chez eBay.

eBay avait pour objectif de développer un moteur de recommandation qui comprenne et apprenne à partir du langage contextuel utilisé par l'acheteur, et sache éliminer rapidement les recommandations produits les moins pertinentes. Plus précisément, dans le cadre du développement de l'application eBay pour l'assistant Google, eBay avait besoin d'un [graphe de connaissances](#) « intelligent » capable d'interpréter le langage naturel afin de stocker et de mémoriser les interactions passées avec les acheteurs et d'en tirer des apprentissages.

eBay a opté pour Neo4j, [base de données de graphes native](#) qui intègre des modèles de probabilités, afin de mieux comprendre les scénarios conversationnels d'achat. Le graphe de Neo4j inclut et utilise à la fois le catalogue de produits et les attributs des interactions avec les acheteurs pendant leur recherche d'articles.

Par ailleurs, le déploiement du robot conversationnel auprès des clients nécessitait d'opérer à l'échelle du Web avec un haut niveau de résilience et de disponibilité, des réponses prévisibles en quelques millisecondes et le soutien d'experts en graphes ayant l'expérience dans ce type de déploiement. Neo4j permet un clustering hautement disponible et assure une performance de lecture et d'écriture exceptionnelle. Même quand le nombre de nœuds atteint des millions, l'application reste très réactive aux requêtes des utilisateurs.

L'application finale comprend la base de données de graphes de Neo4j et des algorithmes de compréhension du langage naturel qui non seulement comprennent le texte, mais tiennent aussi compte des intentions dans l'orthographe et la grammaire pour analyser le sens et le contexte des conversations.

Actuellement, l'application fonctionne dans des conteneurs Docker hébergés dans le cloud. À l'avenir, l'équipe d'eBay entend déployer des robots conversationnels sur diverses plateformes par le biais de plugins tels que Slack et Microsoft.

Conclusion

Les moteurs de recommandation en temps réel permettent aux entreprises de se différencier qu'elles opèrent dans le domaine de la vente au détail, la logistique, le recrutement, les médias, l'analyse des sentiments, la recherche ou la gestion des connaissances. Stocker et interroger des données de recommandation à l'aide d'une base de données de graphes permet aux applications de [fournir des résultats pertinents en temps réel plutôt que des données pré-calculées souvent obsolètes](#).

Alors que les consommateurs ont des attentes de plus en plus élevées (et que leur patience a tendance à diminuer), fournir des recommandations pertinentes en temps réel constitue un avantage concurrentiel plus fort que jamais.

Cas d'utilisation n°3 : Graphes de connaissances



Avec la recherche traditionnelle par mots clés, les résultats sont souvent aléatoires, dilués et de mauvaise qualité. Il n'est pas vraiment possible de poser des questions précises, ni d'obtenir les informations les plus pertinentes et les plus significatives.

De même, [les bases de données relationnelles manquent de souplesse](#) notamment lorsqu'il s'agit d'ajouter de nouveaux types de contenus ou faire des modifications structurelles. Dans ces cas-là, il est nécessaire retravailler les modèles relationnels.

Pourquoi utiliser la technologie des graphes pour construire un graphe de connaissances ?

L'approche standard d'une recherche s'avère un processus fastidieux qui consiste à redéfinir de façon répétitive les termes de la recherche jusqu'à tomber sur quelque chose de pertinent. Ce processus est limité, d'autant plus que les utilisateurs s'habituent à des [recherches contextuelles qui comprennent leurs intentions](#).

Il faut aussi prendre en compte les évolutions à venir :

- La recherche vocale permet de poser des questions aux appareils connectés et d'obtenir les réponses recherchées.
- La fonction de recherche se généralise et va apparaître dans des applications émergentes et davantage d'appareils connectés.
- La recherche payante va nécessiter une connaissance approfondie des produits et un pouvoir d'achat plus important.

Pour améliorer les capacités de recherche d'une entreprise, [un graphe de connaissances détenant des capacités de recherche basées sur les données de graphes](#) est requis pour ne procurer que les résultats pertinents et contextualisés.

VOLUMES DE DONNÉES TOUJOURS PLUS VASTES

Les ensembles de données sont en augmentation constante. Sans un stockage efficace des données capable de gérer ces volumes croissants, les recherches restent limitées.

REQUÊTES DE RECHERCHE PRÉCISES

Toute solution de recherche qui n'est pas en mesure d'approfondir le contexte en identifiant l'historique de l'utilisateur, ses intentions ainsi que le contexte, ne peut pas non plus procurer un avantage concurrentiel.

REQUÊTES DE RECHERCHE SANS AMBIGUÏTÉ

Plus les volumes de données augmentent, plus les résultats de recherche qui se recourent sans aucune pertinence s'accumulent. La qualité de la recherche a tendance à diminuer plus le nombre de résultats de recherche augmente.

Exemple : Centre allemand de recherche sur le diabète

Le [Centre allemand de recherche sur le diabète \(DZD\)](#), fondé en 2009, rassemble des experts de toute l'Allemagne pour développer des mesures efficaces de prévention et de traitement du diabète. Afin de mieux comprendre les causes du diabète, les chercheurs du DZD examinent la maladie sous tous les angles. Ils accumulent ainsi de larges volumes de données et travaillent à établir une base de données de référence pour consolider toutes ces informations.

Le DZD avait besoin de créer un ensemble uniforme de données pour permettre l'accès à toutes les données collectées. Les métadonnées devaient être normalisées et intégrées dans un modèle de base de données de graphes pour comparer les résultats des mesures et éliminer les lacunes ou les redondances inutiles.

À la recherche d'un outil adapté, le Dr. Alexander Jarasch, à la tête du management de la bio-informatique et des données du centre, a choisi Neo4j. Il a ainsi mis en œuvre un nouvel outil interne nommé DZDconnect où la base de données de graphes de Neo4j s'appuie sur les différentes bases de données relationnelles et relie ainsi les différents systèmes et silos de données du DZD.

Grâce à la grande évolutivité et aux performances de Neo4j, les possibilités d'intégration de données sont désormais illimitées. À l'avenir, les données de mesure seront automatiquement transmises à l'outil DZDconnect. Cela signifie que d'autres chercheurs peuvent utiliser ces informations dans le cadre de leur travail sans avoir à réitérer les mesures déjà prises ni gâcher de précieux échantillons. Pour les chercheurs, la base de données de graphes de Neo4j peut répondre à des questions riches et complexes.

« Avec la base de données de graphes de Neo4j, nous avons pu combiner et interroger des données dispersées, » indique le Dr. Jarash. « Même si les données n'ont été intégrées qu'en partie, les requêtes ont déjà révélé des relations intéressantes que nos chercheurs vont pouvoir étudier plus en profondeur. »

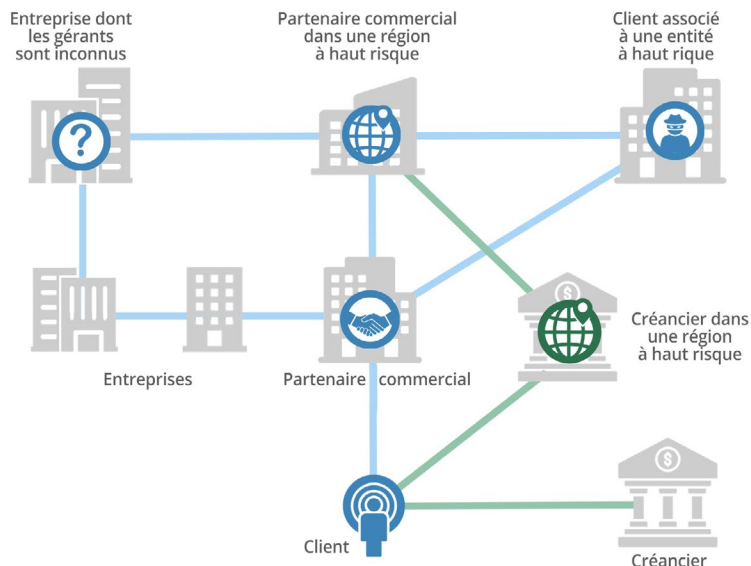
Le Dr. Jarash indique également que la prochaine étape consiste à examiner comment les données humaines des études cliniques peuvent être complétées avec les données hautement standardisées provenant des modèles animaux (par exemple des souris) pour trouver des points communs et autres informations. Les divergences, les modèles et les relations sont faciles à identifier avec une base de données de graphes.

Conclusion

Lorsque les utilisateurs obtiennent les résultats de recherche les plus pertinents (voire les plus inattendus), leur engagement et leur satisfaction augmentent proportionnellement. Or, sans le contexte des relations et des métadonnées, les recherches ne sont pas en mesure de donner les résultats précis qu'attend l'utilisateur.

Avec des volumes de données à la hausse, un graphe de connaissances est nécessaire pour prendre en compte les relations inhérentes à l'ensemble de données. En utilisant un graphe de connaissances, l'accès aux informations s'améliore et les utilisateurs peuvent trouver le produit, le service ou l'actif numérique qu'ils recherchent.

Cas d'utilisation n°4 : Lutte contre le blanchiment d'argent



Aujourd'hui, les modèles de [lutte contre le blanchiment d'argent](#) sont plus élaborés et opèrent souvent de façon indirecte pour détourner et piéger les activités douteuses. Toutefois, les technologies traditionnelles ne sont pas conçues pour relier les points entre les nombreuses étapes intermédiaires de ce processus.

En général, les inspecteurs passent un temps considérable à examiner des larges ensembles de données, souvent plusieurs mois, alors que chaque jour, de nouvelles transactions s'accumulent.

Pourquoi utiliser la technologie des graphes pour lutter contre le blanchiment d'argent ?

Les technologies traditionnelles ne sont pas conçues pour relier les points, si bien que la détection des mécanismes de blanchiment d'argent demande un effort considérable et laborieux. Les équipes d'investigation sont surchargées par l'examen manuel de grandes masses de données.

Entre aussi en jeu la grande variété de tactiques de blanchiment d'argent aujourd'hui à l'œuvre :

- Structuration (alias le smurfing)
- Contrebande d'argent liquide
- Activités manipulant beaucoup de liquidités
- Blanchiment par transactions commerciales
- Sociétés écran et trusts
- Opérations circulaires (round-tripping)
- Extraction de données bancaires
- Casinos and autres jeux d'argent
- Immobilier
- Salaires en liquide
- Activités d'assurance-vie

FAUX POSITIFS

Les alertes relatives aux faux positifs ne cessent de proliférer ce qui érode la confiance dans l'exactitude des données et entraîne des pertes considérables en efforts gaspillés pour le secteur financier.

CONTEXTE RÉGLEMENTAIRE

La réglementation appelle à davantage de supervision et d'analyse. La transparence dans la communication à propos des données est stratégique pour éviter des sanctions pour non-conformité et des poursuites judiciaires.

SOLUTIONS TRADITIONNELLES

Avec le blanchiment d'argent, tout est question de temps. Les processus actuels de lutte contre le blanchiment freinent la productivité et réduisent l'efficacité.

Avec de vastes volumes de données en constante augmentation, une solution technologique capable de s'adapter à l'envergure et à la variété des relations dans les ensembles de données s'impose.

[Les bases de données de graphes, capables d'identifier les relations entre données complexes avec une performance de requête en temps réel](#), offrent une arme puissante contre les mécanismes de blanchiment d'argent et de détournement de fonds.

Exemple : Service de transfert de fonds

Le responsable de la conformité d'un service de transfert de fonds avait besoin d'un outil qui permette aux enquêteurs sur le blanchiment d'argent de repérer les tendances inhabituelles au sein de centaines de milliers de transactions quotidiennes.

Ce service de transfert de fonds, par lequel transitent près de 600 milliards de dollars par an, cherchait un moyen de détecter les activités dites de « smurfing » destinées à disperser de grandes sommes d'argent sale dans un réseau de bénéficiaires.

À juste titre, le secteur des transferts de fonds est strictement réglementé et cette entreprise répond aux obligations de lutte contre le blanchiment d'argent de chaque pays dans lequel elle opère (150 pays dans le monde).

« Nous cherchons constamment à améliorer nos systèmes et nos processus » explique le responsable de la conformité. « Nous avons étudié la technologie de mapping avec la FIU (plateforme des cellules de renseignement financier européennes) et nous avons tout de suite cerné le potentiel de ce type d'outil pour notre activité. »

Quand nous avons découvert Neo4j, nous avons compris qu'il s'agissait du [meilleur outil notamment grâce à son mode d'inspection dynamique des données](#), » poursuit-il. « Nos outils [traditionnels] étant tous basés sur la technologie SQL, il est impossible de suivre l'approche dynamique que Neo4j nous propose dans ce processus d'investigation. »

En quelques clics, Neo4j procède à un traitement dynamique et en temps réel des larges volumes de données de cette entreprise de transfert de fonds et renforce ainsi son efficacité en termes de conformité et de processus d'investigation.

Un cas de figure majeur impliquait plusieurs pays répartis sur deux continents. Le responsable de la conformité explique : « Neo4j nous a permis de visualiser l'ensemble du réseau. Le facteur critique de doute résidait dans l'itinéraire — et il sautait aux yeux en le visualisant sous la forme d'un graphe avec Neo4j. »

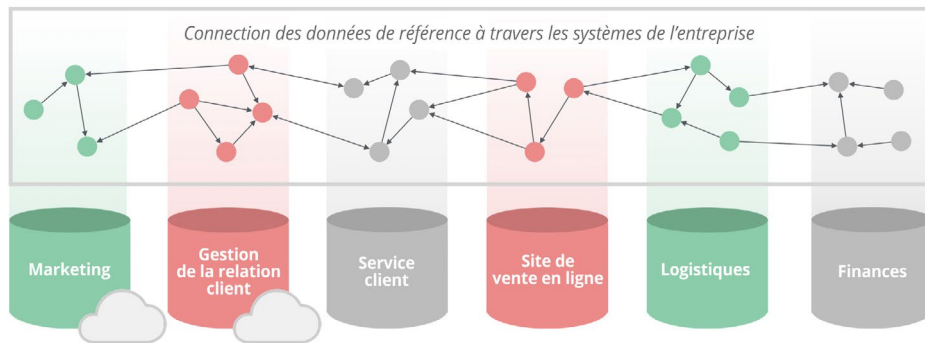
Depuis qu'elle a entièrement déployé Neo4j, l'entreprise est en mesure de traquer les activités criminelles 20 fois plus vite qu'avec ses outils précédents, et a également défini les critères de référence pour la conformité des enquêtes dans son secteur.

Conclusion

Avec une base de données de graphes, le niveau de détection est en constante amélioration notamment grâce à l'ajout de nouvelles sources et nouveaux types de données sans avoir à réécrire le [modèle de données](#). Grâce à des fonctions intégrées de haute disponibilité, l'utilisateur a la garantie que les données sont toujours disponibles pour son moteur stratégique de lutte contre le blanchiment d'argent.

Contrairement aux bases de données relationnelles, un graphe stocke les données interconnectées et facilite la détection des activités frauduleuses quelles que soient la profondeur et la configuration des données.

Cas d'utilisation n°5 : Gestion des données de référence



Les données de référence sont indispensables à l'entreprise, notamment les données concernant :

- Les utilisateurs
- Les clients
- Les produits
- Les comptes
- Les partenaires
- Les sites
- Les unités commerciales

De nombreuses applications métier utilisent des données de référence. Or, ces données sont souvent stockées à différents emplacements, sous divers formats et avec des niveaux de qualité variables. Elles comprennent également des chevauchements et redondances multiples. La gestion des données de référence est la pratique qui consiste à identifier, nettoyer, stocker et surtout gérer ces données.

Les pratiques d'excellence dans la gestion des données de référence s'échelonnent de la fusion de toutes les données de référence en un seul et même emplacement, à la gestion des données pour un accès facilité à partir d'un service ou d'une application unique. Dans les deux cas (ou pour une solution hybride), les architectes des données de l'entreprise ont besoin d'un [modèle de données qui puisse fournir des structures ad hoc](#) capables d'évoluer à mesure que les besoins de l'entreprise changent. Les bases de données de graphes sont parfaitement adaptées à ce modèle de données en constante évolution.

Pourquoi utiliser la technologie des graphes pour la gestion des données de référence ?

Les données de référence étant hautement connectées et partagées, les systèmes de gestion de données de référence mal conçus nuisent à la flexibilité de l'entreprise de telle manière que tous les services de l'entreprise en pâtissent. La plupart des systèmes de gestion de données de référence traditionnels s'appuient sur une base de données relationnelle qui n'est pas optimisée pour traiter les relations ou être réactive aux changements.

Les connexions et les relations entre les données présentes dans des ensembles de données de référence sont essentielles pour obtenir un avantage concurrentiel à mesure que l'analyse métier évolue. Les bases de données de graphes sont idéales pour modéliser, stocker et interroger les hiérarchies, les métadonnées et les relations entre les données de référence.

ENSEMBLES DE DONNÉES COMPLEXES ET HIÉRARCHISÉS

Les données de référence sont étroitement hiérarchisées avec des relations descendantes, latérales et diagonales. Gérer de tels modèles de données avec une base de données relationnelle aboutit à un code complexe et peu maniable, dont l'exécution est lente, l'élaboration coûteuse et la mise à jour chronophage.

PERFORMANCE DE STOCKAGE ET DE REQUÊTE EN TEMPS RÉEL

Le magasin de données de référence doit fournir des données et s'intégrer avec de nombreuses applications au sein de l'entreprise, parfois en temps réel. Mais traverser un ensemble de données complexe et hautement interconnecté pour fournir des informations en temps réel est un défi de taille.

STRUCTURE DYNAMIQUE

Les données de référence sont de nature dynamique et il est difficile pour les développeurs de concevoir des systèmes capables de suivre leurs évolutions.

En comparaison avec la mise en place d'une solution relationnelle, grâce à une base de données de graphes, les données de référence sont bien plus faciles à modéliser et demandent beaucoup moins de ressources (modélisateurs, architectes, administrateurs de bases de données et développeurs). Ainsi, il n'est pas nécessaire de migrer toutes les données de référence vers un seul et même emplacement. Les relations orientées graphe connectent les données réparties en silos avec les systèmes de gestion de la relation client, les systèmes d'inventaire, de comptabilité et de point de vente pour apporter une vue cohérente des données de l'entreprise.

Exemple : Airbnb

Dans une entreprise [aussi étendue et complexe qu'Airbnb](#), un écosystème de données internes et externes en constante évolution finit par devenir ingérable et restrictif, en particulier, quand les données sont dispersées à travers différentes plateformes.

John Bodley, ingénieur logiciel chez Airbnb, a reconnu que les données de l'entreprise étaient excessivement cloisonnées, inaccessibles ou dépourvues de contexte. Il a également remarqué que les employés comptaient sur des connaissances tribales pour répondre à leurs questions, ce qui nuisait à la productivité.

« Nous menons régulièrement un sondage auprès de nos employés » relate-t-il, « et l'affirmation 'l'information dont j'ai besoin pour faire mon travail est facile à trouver' a toujours obtenu de très mauvaises évaluations. » L'équipe de John Bodley savait qu'il était nécessaire de démocratiser les données afin que tout employé, quelle que soit son aisance avec les données, ait les moyens de trouver les ressources recherchées, en étant sûr de la pertinence et de la fiabilité des résultats.

Son équipe a commencé à développer Dataportal, un espace de données intégré en libre-service qui présente une vue contextuelle et holistique des données d'Airbnb. Cet espace permet aux employés de naviguer facilement et rapidement dans leur environnement de données lorsqu'ils ont besoin d'y accéder ou de trouver les réponses nécessaires à leur travail quotidien.

John Bodley et son équipe ont vite compris qu'un graphe était le meilleur format pour représenter leur écosystème de données. Cette observation les a conduits à la base de données de graphes de Neo4j qui offre le moyen le plus rapide de faire des recherches à travers des millions de données connectées en quelques secondes. De plus, Neo4j s'intègre avec tous les langages de programmation de prédilection d'Airbnb tout en leur permettant d'enrichir les classements de recherche grâce à la topologie des graphes.

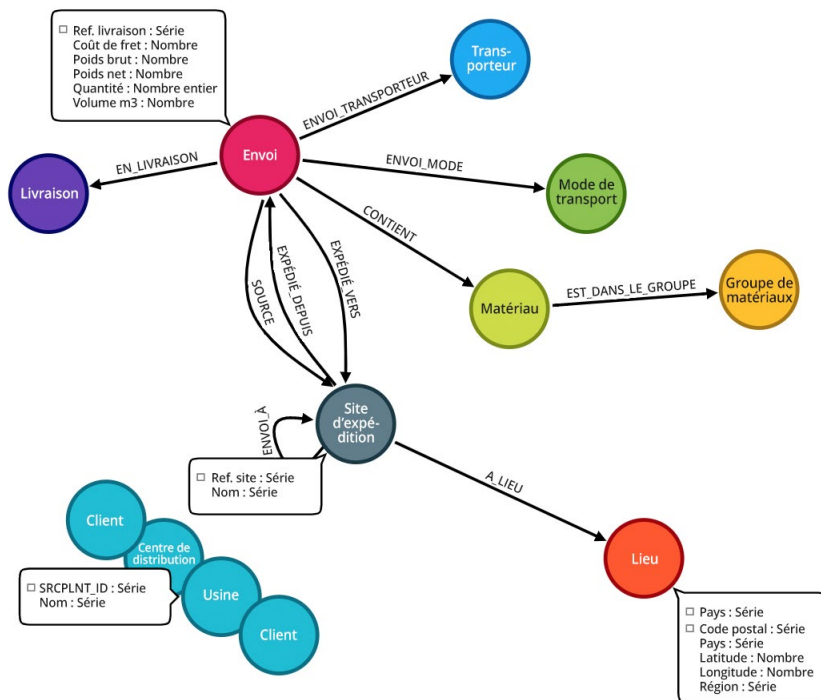
Avec Neo4j, Airbnb a pu fédérer l'ensemble de son écosystème de données et le rendre consultable, pertinent et fiable. Au lieu de travailler de façon dispersée en misant sur des connaissances tribales, le personnel dispose désormais de Dataportal comme ressource unique pour trouver toutes les données pertinentes, essentielles à la performance quotidienne.

Conclusion

Les meilleures décisions métier ne reposent pas sur des données obsolètes réparties en silos mais bien sur des données de référence et leurs relations, disponibles en temps réel.

Les bases de données de graphes sont conçues pour prendre en charge les relations entre les données. Grâce à une modélisation et une interrogation plus efficaces, [l'organisation des données de référence dans un graphe](#) apporte des réponses pertinentes plus rapidement et avec plus de flexibilité que jamais auparavant.

Cas d'utilisation n°6 : Gestion de la chaîne d'approvisionnement



Par nature, la gestion de la chaîne d'approvisionnement est dynamique. Elle inclut de nombreux composants mouvants et peut subir des goulets d'étranglement à chacun de ses maillons. Or, les bases de données traditionnelles ne sont pas capables de réaliser un traitement précis des informations en temps réel à partir des volumes de données détaillées qu'elles génèrent.

Aujourd'hui, les « big data » pilotent de nombreux secteurs, mais la gestion de la chaîne d'approvisionnement fait figure de cas d'usage idéal concernant la visualisation des données, l'analytique prédictive et le contrôle de la productivité. Les données vont continuer à proliférer, si bien que les suivre et les maîtriser va créer des difficultés grandissantes.

Pourquoi utiliser la technologie des graphes dans le cadre de la gestion de la chaîne d'approvisionnement ?

Les données de référence sont hautement partagées au sein de l'entreprise. Ce phénomène a tendance à provoquer une perte d'agilité qui se répercute dans tous les départements. Les architectures de données actuelles s'attachent à obtenir des données qui correspondent à une définition unique de la vérité. Mais bien souvent, il apparaît que cela ne fonctionne pas sur le long terme.

Les segments de données de la gestion de la chaîne d'approvisionnement

- Distribution des ressources du réseau pour répondre à la demande
- Planification d'itinéraires en fonction du temps et des coûts de transport
- Gestion de la relation client

ITINÉRAIRES DE LIVRAISON, LE DERNIER KILOMÈTRE

Une visibilité renforcée des problèmes, efforts et coûts qui pourraient entraver la dernière étape du processus de transport est nécessaire.

TRANSPARENCE ET FIABILITÉ

Se contenter de calculer les coûts et les délais de livraison pour chaque nouveau client n'offre d'intérêt pour aucune partie.

SIMPLIFICATION DES RÉSEAUX DE DISTRIBUTION

Avec un réseau d'entrepôts, d'usines et de centres de distribution, les interrelations présentent de nombreux défis.

PRÉVISION DE LA DEMANDE FUTURE

La demande du marché est versatile et sans une solution de « big data », il n'y a pas de réelle visibilité sur les clients, les fournisseurs, les capteurs, ni les facteurs externes.

- Activités marketing ciblant de nouveaux clients et sources de revenus
- Analyse des risques d'accidents potentiels et d'événements inhabituels
- Travail de maintenance, rénovation et achat de fournitures
- Gestion des centres de distribution et des entrepôts pour une allocation optimisée des capacités et de l'espace disponible

Exemple : Transparency-One

Aujourd'hui, les chaînes d'approvisionnement sont très étendues et diversifiées, si bien qu'elles deviennent un terrain propice aux risques. C'est pourquoi la transparence devient à la fois cruciale et complexe.

Face à ce constat, [Transparency-One a développé une plateforme](#) qui permet aux fabricants et aux marques de mieux comprendre leurs chaînes d'approvisionnement, de les superviser, de les analyser et d'y lancer des recherches afin de partager des données significatives concernant les sites de production et les produits.

À partir de 2013, il était nécessaire d'étendre cette cartographie pour inclure des informations plus détaillées pour englober tous les éléments de la chaîne d'approvisionnement (produits, cartographie de la chaîne, etc.)

Chris Morrison, PDG de Transparency-One, indique : « Le défi était d'autant plus grand que ce domaine était tout nouveau pour nos clients et qu'aucune solution du marché ne proposait une transparence totale. »

Pour développer cette solution, Transparency-One s'est d'abord tournée vers une solution de type base de données relationnelle SQL. Rapidement, il est apparu que le volume et la structure des informations à traiter allaient avoir une incidence majeure sur la performance et causer des problèmes considérables.

L'équipe de Transparency-One a s'est donc tournée vers les bases de données de graphes.

« Une fois que nous avons décidé d'utiliser une base de données de graphes, nous nous sommes intéressés aux bases de données utilisées par les principaux acteurs du marché » explique Chris Morrison. « La réponse était évidente et sans appel : la réputation de Neo4j, le leader mondial du domaine, n'était plus à faire. »

Transparency-One a choisi Neo4j, car seule une base de données de graphes pouvait répondre aux contraintes de sa plateforme. Parmi les facteurs décisifs figure le fait que Neo4j peut gérer de vastes volumes de données et s'avère être la base de données de sa catégorie la plus utilisée au niveau mondial.

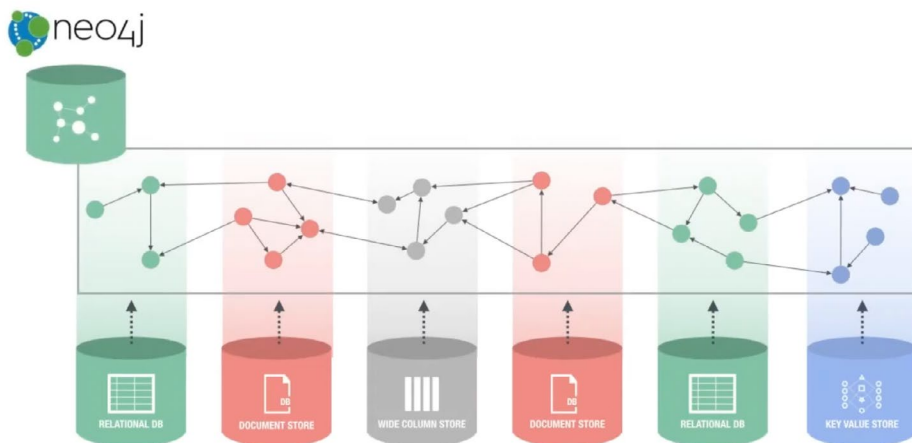
Le projet a rapidement été mis en place et un prototype a été développé en moins de trois mois. Depuis, Transparency-One a étendu la solution en y ajoutant de nouveaux modules et plusieurs entreprises clientes l'utilisent déjà.

Conclusion

Les systèmes de base de données relationnelles ont leurs avantages, mais ils coûtent cher et manquent d'agilité et de vélocité quand il s'agit de traverser un réseau de données de référence.

Avec une base de données de graphes, il est possible d'effectuer un stockage qui procure [une vue à 360 degrés pour gérer et exploiter efficacement les relations entre les données de la chaîne d'approvisionnement](#). De plus, le modèle de graphe évolue facilement au fur et à mesure que les données changent, ce qui apporte une meilleure performance et une grande évolutivité. Ceci permet de prendre des décisions en temps réel, indépendamment de l'ampleur de l'ensemble de données.

Cas d'utilisation n°7 : Renforcer la gestion des opérations informatiques et de réseau



Données stockées en silos dispersés

Par nature, les réseaux sont des graphes. C'est pourquoi les bases de données de graphes conviennent parfaitement pour [la modélisation, le stockage et l'interrogation des données relatives aux opérations informatiques et de réseau](#).

Aujourd'hui, les bases de données de graphes sont utilisées efficacement dans les domaines des télécommunications, de la gestion de réseau, de l'analyse d'impact, de la gestion des plateformes cloud, des centres de données et des ressources informatiques.

Dans chacun de ces domaines, les bases de données de graphes stockent des informations de configuration afin d'alerter les opérateurs en temps réel à propos des défauts de conception potentiels et des défaillances partagées au sein de l'infrastructure. Ceci permet un raccourcissement des temps d'analyse et une résolution plus rapide des problèmes, qui passent de quelques heures à quelques secondes.

Pourquoi utiliser la technologie des graphes pour la gestion des opérations informatiques et de réseau ?

Comme avec les données de référence, une base de données de graphes permet de regrouper des informations issues de systèmes d'inventaire disparates offrant ainsi une vision unifiée du réseau et de ses utilisateurs ; des plus petits éléments du réseau jusqu'aux applications, services et clients qui les utilisent.

Une représentation par les graphes d'un réseau permet aux responsables informatiques de cataloguer les ressources, de visualiser leur déploiement et d'identifier les dépendances qui existent entre les deux. La structure connectée du graphe permet aux gestionnaires de réseau d'effectuer des analyses d'impact sophistiquées, en répondant à des questions telles que :

- De quelles portions du réseau (applications, services, machines virtuelles, machines physiques, centres de données, routeurs, commutateurs et fibre optique) les clients dépendent-ils ? (Analyse descendante)

DÉPANNAGE

Modification majeure du réseau, renforcement de la sécurité des accès ou optimisation de l'infrastructure applicative, quel que soit le scénario, les interdépendances sur le plan physique et humain sont d'une complexité extrême et rendent les dépannages de plus en plus complexes.

ANALYSE D'IMPACT

Les relations entre différents nœuds du réseau ne sont ni purement linéaires, ni purement hiérarchisées, d'où la difficulté à déterminer l'interdépendance entre les composants du réseau. Ces relations deviennent encore plus complexes quand deux systèmes (ou plus) fonctionnent ensemble.

CROISSANCE DES NŒUDS PHYSIQUES ET VIRTUELS

Face à une croissance rapide de la taille des réseaux et du nombre de composants nécessaires à la prise en charge de nouveaux services et périphériques connectés, les services informatiques doivent développer des systèmes qui répondent aux exigences actuelles et futures.

- À l'inverse, quels services et applications et, en définitive, quels clients du réseau seront affectés si un élément spécifique de celui-ci tombe en panne ? (Analyse ascendante)
- Y a-t-il une redondance à travers le réseau pour les clients les plus importants ?

La représentation d'un réseau sous forme de graphes permet également d'enrichir les informations opérationnelles en fonction des corrélations d'événements.

Chaque fois qu'un moteur de corrélation d'événement (tel qu'un [processeur d'événements complexes](#)) déduit un événement complexe à partir d'un flux d'événements de réseau, il évalue l'impact de cet événement par rapport au modèle graphe et déclenche les mesures de compensation ou d'atténuation nécessaires.

Exemple : Telia

Telia Zone, routeur de la société Telia, équipe près de 1,5 million de foyers en Suède. [Telia utilise Neo4j pour connecter tous les domiciles de ses clients](#) et héberge des clusters causaux avec Kubernetes afin de représenter toutes les actions qui ont lieu dans et à travers les routeurs dans un graphe.

Telia a décidé d'étendre les cas d'utilisation que permet son routeur et se sert de ce graphe pour déterminer combien de cas d'utilisation sont possibles. La démarche a d'autant plus d'intérêt que l'entreprise poursuit le développement des fonctionnalités du routeur au-delà d'une connexion élémentaire à Internet.

Certains de ces cas d'utilisation concernent par exemple la connexion de haut-parleurs Sonos à Spotify pour plusieurs utilisateurs dans un même foyer, ou encore la notification par sms du retour à la maison d'un enfant à ses parents. Toutes ces actions créent des instances qui sont suivies à travers la base de données de graphes pour éviter de surcharger les serveurs et identifier les utilisateurs présents dans une même zone au même moment. Ces cas d'utilisation sont possibles grâce aux API de Telia.

« Ce que nous voulons procurer à nos clients, c'est l'identification des utilisateurs. Non seulement les adresses MAC que nous utilisons en interne, mais quelque chose qui a davantage de valeur, » explique Lars Ericsson, responsable de zone et ingénieur chez Telia. « Nous voulons fournir les noms des utilisateurs, les tokens ou tout autre identifiant utilisés, et nous stockons toutes ces informations dans Neo4j. »

Du côté des routeurs, plus de 1,5 milliard de requêtes sont effectuées chaque jour. Grâce au clustering causal dans Neo4j, Telia procède par étapes, au lieu de tout transmettre en une seule fois à tous les foyers.

« Nous ne voulons pas saturer nos serveurs avec le traitement de toutes les requêtes qui peuvent intervenir, » poursuit Lars Ericsson, « nous voulons des serveurs qui traitent les requêtes entrantes en priorité. »

Conclusion

Découvrir et donner un sens à des interdépendances complexes est essentiel pour gérer efficacement les opérations informatiques et de réseau. Qu'il s'agisse d'optimiser une infrastructure de réseaux / d'applications ou de fournir un accès plus sécurisé, ces problèmes impliquent un ensemble complexe d'interdépendances physiques et humaines dont la gestion relève du défi.

Les relations qui existent entre les éléments du réseau et de l'infrastructure sont rarement linéaires ou purement hiérarchiques. Les bases de données de graphes sont conçues pour stocker ces données interconnectées et permettre de [traduire facilement les données informatiques et de réseau en informations exploitables](#).

Cas d'utilisation n°8 : Traçabilité des données

[La modélisation des risques](#) présente de multiples contraintes. Elle implique également que les organisations — en particulier les grandes banques, les fonds spéculatifs et les sociétés d'investissement — suivent les relations entre données à travers un réseau complexe d'investissements, de participations, d'instruments financiers et de données granulaires sur les prix.

Pourquoi utiliser la technologie des graphes en matière de traçabilité des données ?

Les banques et autres organismes financiers réglementés sont tenus de retracer les dépendances entre les données à de nombreux niveaux qui s'entrecoupent avant d'atteindre les sources de données confirmées.

La plupart des systèmes ne peut pas répondre à cette exigence cruciale. Les problématiques de traçabilité des données que soulèvent des exigences strictes comme celles de la norme BCBS 239 demandent bien plus d'agilité et de persistance que ne le proposent les systèmes traditionnels de stockage de données.

En effet, il est nécessaire de [remonter la trace des données par le biais de silos discrets jusqu'à atteindre une source de données confirmée](#). De même, quoique les entités requièrent des identifiants standardisés, les unités opérationnelles utilisent bien souvent leurs propres termes et algorithmes, avec des distinguos parfois au cœur d'une même entité.

La structure et la localisation des données empêchent parfois de recourir à un stockage des données unique et centralisé. Ironiquement, regrouper toutes les données en un seul emplacement peut compliquer davantage la traçabilité des données.

Intégrer ces informations dans un modèle de données unique à l'échelle de l'entreprise nécessite la technologie des graphes. Avec une base de données de graphes, il est possible d'obtenir une source unique de la situation réelle, de requêter et de mettre à jour les relations entre données complexes ou dissimulées en quelques millisecondes.

Exemple : UBS

Sid Hannif, architecte de solutions chez UBS, se démenait avec des centaines de systèmes de données cloisonnés qu'il devait relier pour [offrir une expérience fluide aux utilisateurs grâce à des données mises à jour en temps réel](#).

« Nous sommes censés fournir des données actualisées, précises et complètes, mais l'utilisateur se retrouvait obligé de relier les données par lui-même, » relate-t-il.

UBS opère depuis 1862 et compte environs 66 000 employés. Ce grand acteur de la finance est réparti en plusieurs division qui recoupent la gestion de patrimoine, la banque d'investissement, la gestion d'actifs, un centre d'entreprises et des services de données.

Dans la division des services de données, Sid Hannif et son équipe sont responsables de 12 domaines et sous-ensembles de données, depuis lesquels les données sont diffusées sur plusieurs canaux. Les données de comptes résidaient dans des bases de données relationnelles mais le nombre d'intégrations posait un obstacle majeur et la plateforme de diffusion des données de référence sur les clients d'UBS en pâtissait.

GOVERNANCE ET INFRASTRUCTURE

Les banques doivent pratiquer la gouvernance des données et utiliser des taxonomies de données intégrées, ainsi que des données de référence à l'échelle du groupe qui comprennent les identifiants d'entités, de partenaires, de clients et de comptes.

AGRÉGATION DES DONNÉES SUR LE RISQUE

Les banques doivent générer des données précises, cohérentes et fiables sur les risques. En parallèle, elles doivent maintenir une visibilité totale concernant les sources de données confirmées. Ces ensembles de données doivent répondre à un éventail complet de requêtes émanant de la direction et des régulateurs.

RAPPORT DES RISQUES

Les rapports réglementaires et de gestion doivent représenter les risques de manière précise, adaptée à des fins d'audit, et correspondre à la complexité du modèle des risques et des opérations de la banque.

EXAMEN DE SUPERVISION

Les contrôleurs des risques bancaires sont tenus d'examiner la conformité permanente de leur établissement aux principes de la norme BCBS 239.

L'équipe des services de données voulait que les données de référence soient utilisées par les moteurs de requête pour interroger leur modélisation et leurs liens avec d'autres données, pour ensuite les relier. C'est exactement ce que réalise un graphe de connaissances en tant que sous-composant du système.

UBS a choisi la base de données de graphes de Neo4j pour modéliser ses données.

« La gouvernance des données, notamment leur déplacement d'un endroit à l'autre, et désormais bien plus facile » indique Sid Hannif. « Nous comprenons qui se sert de nos données, sur quels canaux. Plus important encore, nos utilisateurs sont en mesure de comprendre les ensembles de données que nous proposons et quels sont les liens entre eux. »

Conclusion

Utiliser une base de données de graphes à des fins de conformité en matière de protection des données personnelles et de gestion des risques offre plusieurs avantages que n'offrent pas les systèmes traditionnels :

- Traçabilité des facteurs de risques jusqu'à la source d'origine confirmée
- Regroupement de silos de données sur la tarification, la position, la gestion de la trésorerie, entre autres, en un ensemble de données unifié
- Collaboration avec les régulateurs pour visualiser et faire évoluer le modèle de risques basés sur les graphes
- Modification facilitée des modèles de risques pour suivre l'évolution des conditions du marché, les changements organisationnels et les stratégies d'investissement
- Traitement des fusions/acquisitions, des cessions et des réorganisations qui affectent le fonctionnement et les performances historiques et futures des bureaux de négociation

La base de données de graphes de Neo4j prend également en charge les normes terminologiques financières mondiales soutenues par des services professionnels garants de réussite et source d'une nouvelle visibilité sur les mesures de conformité et les opérations quotidiennes.

Cas d'utilisation n°9 : Gestion des identités et des accès

Les solutions de gestion des identités et des accès stockent des informations sur les parties (administrateurs, divisions commerciales, utilisateurs finaux) et les ressources (fichiers, partages, périphériques réseau, produits, accords), ainsi que les règles gouvernant l'accès à ces ressources. Les solutions de gestion des identités et des accès appliquent ces règles pour déterminer qui à l'autorisation de consulter ou manipuler une ressource donnée.

Traditionnellement, la [gestion des identités et des accès](#) était mise en œuvre en utilisant des services de répertoire ou en concevant une solution personnalisée pour une application. Cependant, les structures de répertoires hiérarchiques ne sont pas capables de gérer les structures de dépendances complexes qui existent dans les chaînes d'approvisionnement multipartites distribuées. Les solutions personnalisées utilisant des bases de données traditionnelles pour stocker les données relatives aux identités et aux accès deviennent lentes et peu réactives à mesure que le volume des ensembles de données augmente.

Pourquoi utiliser la technologie des graphes pour la gestion des identités et des accès ?

Une base de données de graphes peut stocker des structures de contrôle d'accès complexes et hautement connectées, englobant des milliards de parties et de ressources. Son modèle de données sophistiqué et variable prend à la fois en charge les structures hiérarchiques et non hiérarchiques, tandis que son modèle de propriété extensible permet de capturer les métadonnées concernant chaque élément du système.

Avec un moteur de requête capable de parcourir des millions de relations par seconde, les recherches d'accès effectuées avec une base de données de graphes, à travers des structures volumineuses et complexes, s'effectuent en quelques millisecondes, et non pas en minutes ou en heures.

Tout comme avec les opérations informatiques et de réseau, une solution de contrôle des accès reposant sur une base de données de graphes permet d'effectuer des requêtes ascendantes et descendantes :

- Quelles ressources (structure de l'entreprise, produits, services, accords et utilisateurs finaux) un administrateur spécifique peut-il gérer ? (Analyse descendante)
- Qui peut modifier les paramètres d'accès d'une ressource donnée ? (Analyse ascendante)
- À quelles ressources un utilisateur final peut-il accéder ?

Les solutions de contrôle d'accès et d'autorisation se basant sur les bases de données de graphes s'appliquent particulièrement aux domaines de la gestion de contenu, des services d'autorisation fédérés, des préférences en termes de réseaux sociaux et des services cloud (SaaS), où, par rapport aux bases de données relationnelles, elles améliorent les performances, qui ne se mesurent alors plus en minutes mais en millisecondes.

DONNÉES D'AUTORISATION INTERCONNECTÉES

Pour vérifier la validité d'une identité et ses autorisations d'accès, un système doit interroger un ensemble de données interconnectées dont le volume et la complexité augmentent en permanence.

PRODUCTIVITÉ ET SATISFACTION CLIENT

À mesure que le nombre d'utilisateurs, de produits et d'autorisations augmente, les systèmes traditionnels n'assurent plus une bonne performance de requête, ce qui se traduit par une diminution de l'expérience des utilisateurs.

STRUCTURE ET ENVIRONNEMENT DYNAMIQUES

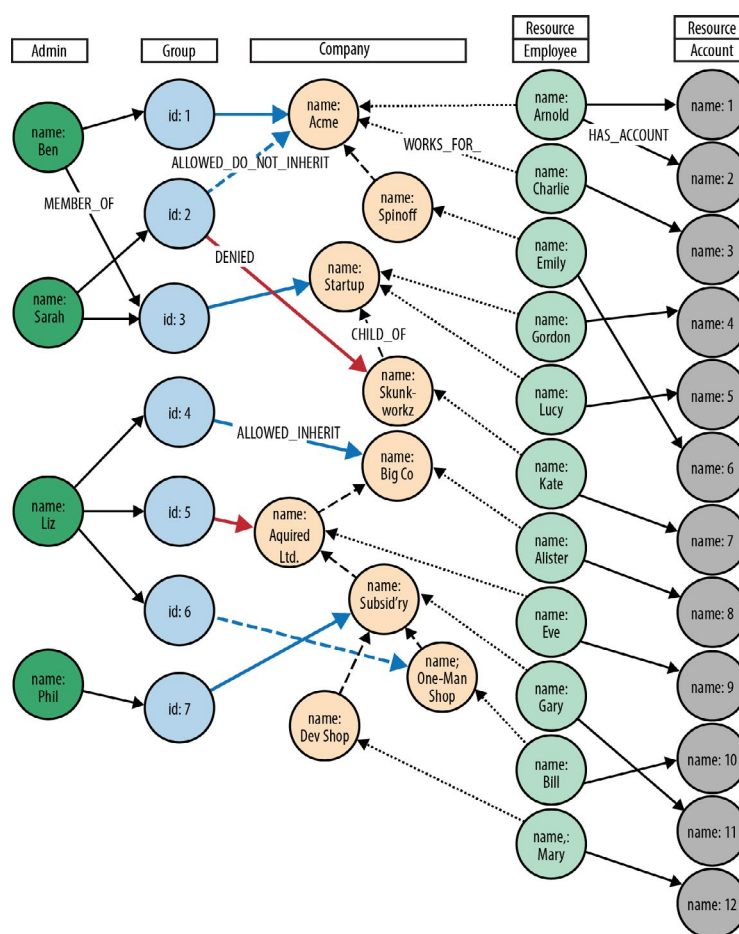
Avec la croissance rapide du volume d'utilisateurs et des métadonnées associées, les applications doivent prendre en compte les besoins de gestion des identités actuels et futures.

Exemple: Telenor Norway

[Telenor Norway](#) est une entreprise internationale de télécommunications. Depuis plusieurs années, elle permet à ses plus gros clients d'utiliser leurs comptes en libre-service. À l'aide d'une application basée sur un navigateur, les administrateurs de chacune des entreprises clientes peuvent ajouter ou supprimer des services au nom de leurs employés.

Pour s'assurer que les utilisateurs et les administrateurs ne voient et ne modifient que les parties de l'entreprise et les services qu'ils ont le droit de gérer, l'application fait appel à un système complexe de gestion des identités et des accès qui attribue des privilèges à des millions d'utilisateurs à travers des dizaines de millions d'instances de produits et de services.

Ci-dessous se trouve un exemple du modèle de données de Telenor :



Légende : échantillon du modèle de données de Telenor Norway présentant son application de gestion des identités et des accès.

Suite à des problèmes de performances et de réactivité, Telenor a décidé de remplacer son système de gestion des identités par une solution de base de données de graphes. Son système initial utilisait une base de données relationnelle qui faisait appel à des jointures pour modéliser des structures organisationnelles et des hiérarchies de produits complexes. En raison de cette structure, leurs requêtes les plus complexes étaient d'une lenteur inacceptable.

En revanche, après avoir implémenté une solution de base de données de graphes, Telenor a obtenu les performances, l'évolutivité et la réactivité nécessaires pour répondre à ses besoins de gestion des identités et des accès, réduisant à quelques millisecondes les requêtes qui auparavant nécessitaient de nombreuses minutes.

Conclusion

Pour une entreprise, gérer de multiples rôles, groupes, produits et autorisations en constante évolution est une tâche qui s'avère de plus en plus complexe. Les bases de données relationnelles ne sont tout simplement pas à la hauteur pour la gestion des identités et des accès car les requêtes sont beaucoup trop lentes et peu réactives.

Avec une base de données de graphes, vous pouvez [suivre facilement et en temps réel les relations entre les identités et les accès](#). Avec une vue interconnectée de vos données, vous disposez de meilleures informations et de contrôles jamais obtenus auparavant.

Cas d'utilisation n°10 : Nomenclatures

L'ère de la transformation numérique n'a pas rendu les nomenclatures obsolètes qui restent bien au cœur des processus de fabrication. Le problème vient plutôt de l'éparpillement des données de fabrication.

Le manque de collaboration en temps réel avec un ensemble dispersé de personnes (fournisseurs, sous-traitants), autour de produits multidisciplinaires (mécaniques, électromécaniques) alliés à de nouveaux flux de travail conception-fabrication, aboutit à des tableurs et des données cloisonnées qui font obstacles aux résultats.

Pourquoi utiliser la technologie des graphes pour la gestion des nomenclatures ?

La plupart des entreprises de fabrication utilisent des applications logicielles tels que des systèmes de gestion client, de gestion du travail, de comptes créditeurs et comptes débiteurs, systèmes de points de vente, etc. En raison de cette approche, les données doivent être stockées et modélisées sous forme de graphes, où un graphe natif stocke [les données de référence interconnectées](#), lesquelles ne sont jamais purement linéaires ou hiérarchisées.

De plus, les nouvelles tendances en matière de nomenclature créent des défis supplémentaires. Les obligations de conformité sont plus exigeantes. Les pièces contrefaites inondent le marché. On assiste à une consolidation entre fabricants de pièces et les cycles de vie des pièces raccourcissent. En parallèle, l'innovation technologique s'avère augmenter le risque de pénurie tout en contribuant à l'obsolescence.

Un modèle de données flexible, basé sur les graphes permet de faire évoluer facilement les données de référence en fonction de l'évolution des besoins.

Exemple : Armée américaine

« L'armée américaine achète et supervise des millions de pièces d'équipement chaque année » indique Preston Hendrickson, responsable technique du projet CALIBRE de l'armée américaine.

Le défi consistait à [collecter et combiner rapidement un volume massif d'informations de nomenclature](#) – chaque pièce et son coût, l'équipement associé, son cycle de vie estimé et le délai moyen attendu avant défaillance.

Forte de plus d'un million de soldats de terrain, de garde et de réserve ainsi que près de 200 000 employés civils, l'armée américaine déploie également une quantité stupéfiante d'équipement – armes légères, fusils et mitrailleuses, chars, camions et véhicules blindés, et des milliers de navires, d'hélicoptères et d'avions.

L'achat et la gestion d'équipement à une telle échelle soulève un défi logistique considérable. Celui-ci implique d'acquérir des millions de pièces pour des centaines de milliers d'armes et de véhicules chaque année et d'entretenir ces composants – parfois dans des territoires lointains et hostiles. Il est important de noter que réaliser cette tâche permet non seulement d'épargner des coûts mais peut aussi sauver des vies.

Pour le système vieillissant de l'armée américaine, basé sur un ordinateur central, il devenait de plus en plus difficile, voire impossible, de gérer un tel niveau de données.

« Grâce à la base de données de graphes de Neo4j, l'armée américaine peut dorénavant [stocker, explorer et visualiser rapidement son patrimoine de données logistiques](#) » déclare Preston Hendrickson. Il remarque également un contraste saisissant avec le précédent système.

FABRICATION DES PRODUITS

La nomenclature doit être maintenue et mise à jour pour suivre le rythme de développement des produits.

NOMENCLATURES MULTI-DOMAINES

La fabrication ne se cantonne pas aux pièces mécaniques mais comprend également de l'électronique, du logiciel, et d'autres composants.

FLUX DE CONCEPTION

Pas de traçabilité et d'audit au niveau mondial, pas de catalogue centralisé pour la gestion des stocks ou le processus de conception sur commande.

Top 10 des cas d'utilisation des bases de données de graphes

« Auparavant, il fallait 60 heures pour charger les données afin que nous puissions comprendre que telle ou telle pièce devait être remplacée, ou pour obtenir une estimation et une analyse des coûts » poursuit-il. « À présent, cela ne prend que sept à huit heures. »

Les pièces sont parfois commandées par plusieurs millions. Avec Neo4j, l'armée américaine anticipe la demande de pièces de rechange et la répartit en plusieurs périodes, au lieu de la traiter par trimestre. « Il en résulte des commandes et une budgétisation optimisées » indique Preston Hendrickson. « Nous obtenons une estimation plus complète du coût total, mais aussi un coût total de possession plus prévisible. »

« Neo4j fait gagner beaucoup de temps aux analystes » remarque Preston Hendrickson. « Dorénavant, tout est dans le graphe et nous sommes en mesure de visualiser des données plus détaillées, auparavant inaperçues. Nous disposons désormais de données originelles et d'une analyse beaucoup plus affinée. Les réponses sont immédiates. En conséquence, la livraison des pièces est plus précise et le délai de commande est beaucoup plus rapide. »

Conclusion

En utilisant la technologie des graphes, les équipes en charge des nomenclatures obtiennent [une vue cohérente et significative de leurs données de référence](#). Elles peuvent identifier les relations entre les personnes, les comptes, les unités opérationnelles, les transactions et d'autres données.

Une vue élargie de la nomenclature permet de tirer parti de relations contextuelles entre les données. Grâce à un modèle de graphes polyvalent, il est facile de faire évoluer le système au fur et à mesure que les sources de données changent. La performance et l'évolutivité permettent de prendre des décisions en temps réel, quelle que soit la taille de l'ensemble de données.

Conclusion

Ces dix cas d'utilisation des bases de données de graphes ne constituent pas une liste exhaustive mais ils mettent en évidence des applications notoires et rentables basées sur la technologie des graphes. D'ailleurs, il existe de nombreux autres cas d'utilisation de la technologie des graphes notamment dans le cadre [des sciences de la vie](#), [des réseaux sociaux](#), [des jeux en ligne](#), [du gouvernement](#), [des sports](#) et mêmes [des organisations à but non-lucratif](#).

Aujourd'hui, les directeurs des systèmes d'information et de la technologie sont de plus en plus contraints de fournir des informations exploitables à partir de leurs données, alors même que les ensembles de données sont de plus en plus volumineux et difficiles à manipuler. Ce dont ils ont besoin est une technologie capable de déterminer les connexions entre les données et d'en tirer des conclusions pertinentes et appropriées.

Les bases de données de graphes sont la solution. Elles permettent à tous les professionnels des données d'exploiter le potentiel des relations entre les données plutôt que de se contenter de point de données individuels. La seule limite dans la manière dont ces relations peuvent être maîtrisées réside dans l'imagination de l'utilisateur.

Ainsi, les bases de données de graphes sont une vague grandissante – et bien plus qu'une mode passagère – dans le monde de l'analytique.

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](#).

Des questions à propos de Neo4j ?

Contactez-nous :
info@neo4j.com
neo4j.com/contact-us