

EXAMENSARBETE A Comparison of Relational and Graph Databases for CRM Systems**STUDENTER** Jan Zubac, Victor Winberg**HANDLEDARE** Per Andersson (LTH), Johan Groth (Lime)**EXAMINATOR** Flavius Gruian (LTH)

Grafdatabas, den nya lagringsmetoden?

POPULÄRVETENSKAPLIG SAMMANFATTNING **Jan Zubac, Victor Winberg**

Det finns flertal fall då grafdatabaser överträffar traditionella relationsdatabasers prestanda. Är grafdatabaser potentiellt den nya lagringsmetoden vid stor mängd sammankopplad data? Vilka villkor krävs? För att sätta det på prov har vi skapat en testmiljö med en stor mängd relationsbunden data.

På sistone har grafdatabasers popularitet och därmed även användningen av dessa databastyper ökat avsevärt. Några av de hemsidor med störst besökarantal såsom Facebook och LinkedIn använder sig i stor utsträckning av grafdatabaser för att lagra sin data. Anledningen till denna ökning är det faktum att dagens applikationer lagrar stora mängder data som är sammankopplad med ett flertal relationer. Traditionella databaser, som lagrar data på ett sätt som kan liknas vid ett Excel-ark med stora tabeller, sägs ha svårt att hantera de nya kraven som ställs på databaser i dagens teknikdrivna samhälle. Stämmer detta eller är de traditionella tabellutformade databaserna trots allt bättre?

Grafdatabaser använder sig av grafstrukturer för att lagra data. Dessa strukturer består av noder som lagrar data och anslutningar som kopplar ihop datan.

I vårt examensarbete har vi undersökt om prestandan i ett kundhanteringssystem, även kallat CRM-system, kan förbättras genom att använda sig av grafdatabasen Neo4j. Dessa typer av system genomför komplicerade operationer som ställer höga krav på databaserna som hämtar datan. Dessa komplicerade operationer gör att det kan ta uppemot en halv minut innan den efterfrågade

datan kan visas för användaren, vilket är problematiskt i vardagligt bruk. Därmed var det intressant att jämföra två helt olika databastyper för att se vilken som gjorde bäst ifrån sig och om det därmed var intressant att övergå till att använda sig av grafdatabaser i CRM-system.

Enligt förespråkare för grafdatabaser är de snabbare än traditionella databaser i ett flertal fall. Det CRM-system som vi har gjort studien på, Lime CRM, använder sig i dagsläget av databassystemet Microsoft SQL Server och den lagrade datan uppfyller många av faktorerna som antyder att grafdatabaser bör briljera i systemet. Men resultatet visar något helt annat.

De traditionella relationsdatabaserna presterade bättre i de flesta användningsfallen av CRM-system, trots de påstådda motgångarna som denna databastyp sägs uppleva vid hantering av stora mängder väldigt sammankopplad data. En av anledningarna till detta är att Limes system inte uppnår en tillräckligt stor databasstorlek för att grafer ska vara fördelaktiga att använda.

Studiens resultat kan användas för att mer specifikt undersöka när grafdatabaser överträffar traditionella databaser, så att företag, specifikt inom kundhantering, vet vilken databastyp som hade varit bäst för deras system.