

Performanz Evaluation von Graphdatenbanksystemen versus MySQL am Beispiel von Kookkurrenzgraphen

Martin Junghanns, Sascha Ludwig & Robert Schulze

13. Januar 2012

Inhalt

1 Einleitung

Aufgabenstellung

- Evaluation verschiedener Graphdatenbanken
- Gegenüberstellung mit einer relationalen Datenbank
- Datenbasis des Wortschatz Leipzig Projekts
- Implementierung eines Frameworks zur Durchführung von Benchmarks in Java
- Auswertung und Visualisierung der Messergebnisse

Graphdatenbanken — allgemein

- Datenmodellierung als Graph
- Vermeidung von Verbundoperation
- Schemafreiheit
- Anwendung: Soziale Netzwerke, Recommendation, Transportnetze

Graphdatenbanken — Auswahl

Neo4j gute Skalierbarkeit, Traversal-Framework, geringe Speicheranforderungen

OrientDB Feature-Vielfalt: SQL-Abfragesprache; Embedded oder als Server

DEX Backend in C++ mit Java-API; gilt als performanteste Graphdatenbank auf dem Markt

Schemavergleich

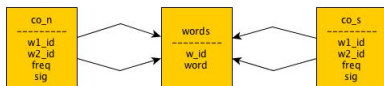


Abbildung: default

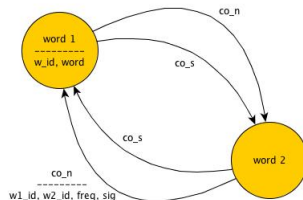


Abbildung: default

Testdaten

■ Kookkurrenzgraph des Wortschatz Leipzig

Datenbank	Knoten	Kanten	
		co_s	co_n
	words		
deu_news_2009_10K	39334	22062	5887
deu_news_2009_100K	191041	297422	69297
deu_news_2009_300K	390178	953218	196122
deu_news_2009_1M	830641	3349314	584402
deu_news_2009_3M	1623621	10596746	1546638
deu_news_2009_10M	3309332	37393614	4412321

Query1 erklären

-
-
-

Query2 erklären

-
-
-

Auswertung

-
-
-

-
-
-

Ende