Ten. Million. Questions. Let's celebrate all we've done together.

Stack Overflow is a question and answer site for professional and enthusiast programmers. It's 100% free.

Take the 2-minute tour

An Example Showing the Necessity of Relationship Type Index and Related Execution Plan Optimization



Suppose I have a large knowledge base with many relationship types, e.g., hasChild, livesIn, locatedIn, capitalOf, largestCityOf...

The number of capicalOf relationships is relatively small (say, one hundred) compared to that of all nodes and other types of relationships.

I want to fetch any capital which is also the largest city in their country by the following query:

MATCH city-[:capitalOf]->country, city-[:largestCityOf]->country RETURN city

Apparently it would be wise to take the capitalOf type as clue, scan all 100 relationship with this type and refine by [:largestCityOf]. However the current execution plan engine of neo4j would do an AllNodesScan and Expand. Why not consider add an "RelationshipByTypeScan" operator into the current query optimization engine, like what NodeByLabelScan does?

I know that I can transform relationship types to relationship properties, index it using the legacy index and manually indicate

START r=relationship:rels(rtype = "capital0f")

to tell neo4j how to make it efficient. But for a more complicated pattern query with many relationship types but no node id/label/property to start from, it is clearly a duty of the optimization engine to decide which relationship type to start with.

I saw many questions asking the same problem but getting answers like "negative... a query TYPICALLY starts from nodes... ". I just want to use the above typical scenario to ask why once

Thanks!

neo4j

share improve this question

asked Apr 8 at 5:08 inzaghi250

add a comment

1 Answer

active oldest

A relationship is local to its start and end node - there is no global relationship dictionary. An operation like "give me globally all relationships of type x" is therefore an expensive operation you need to go through all nodes and collect matching relationships.



There are 2 ways to deal with this:

1) use a manual index on relationships as you've sketched

2) assign labels to your nodes. Assume all the country nodes have a Country label. Your can rewrite your query:

MATCH (city)-[:capital0f]->(country:Country), (city)-[:largestCity0f]->(country) RET

The AllModesScan is now a ModeRyl shallscan. The query grahe all countries and matches to the

asked 4 months ago

viewed 30 times

active 4 months ago

Blog

Why Stack Overflow is a Good Workplace for Women



Looking for a job?

- Implementation Consultant FAST Enterprises Olympia, WA / relocation vb.net sql
- Senior DevOps Engineer CrystalCommerce Mountlake Terrace, WA / remote ruby-on-rails devops

Related

- 2 Neo4j index creation on a database
- o neo4j node property type suggestion, indexes or parent object relationship?
- 2 Error Updating Nodes in Neo4J via Rest
- 3 Neo4j performance is it really fast?
- 1 Difference between START n = node(*) and MATCH (n)
- 0 How do I chose neo4j settings based on the properties of my graph.db?
- 0 Linkurious console query not showing relationship type(name)
- 2 Nesting cypher queries with a collection

cities. Since every country does have one capital and one largest city this is efficient and scales independently of the rest of your graph.

If you put all relationships into one index and try to grab to ~100 capitalOf relationships that operation scales logarithmically with the *total* number of relationships in your graph.

share improve this answer



Thanks for the comments, @stefan - inzaghi250 Apr 11 at 6:54

add a comment

Your Answer



Sign up or log in



Post as a guest

Name	
Email	
required, but never shown	

Post Your Answer

By posting your answer, you agree to the privacy policy and terms of service.

Not the answer you're looking for? Browse other questions tagged neo4j or ask your own question.

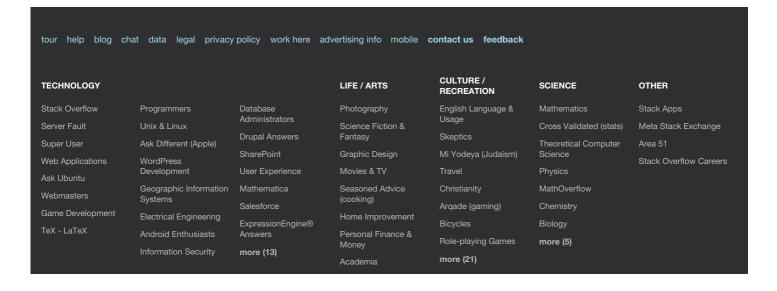
- How to create index on relationship type (neo4j)
- performance of time filters in neo4j where

Hot Network Questions

- Preventing a Print expression from generating a new cell each time it executed
- What's the philosophy behind delaying writing data to disk
- Why are carpenter's pencils flat?
- How to show pagination "go to page" text box?
- Remove a language from the "Preferred language order"
- Why don't decoupling/bypass capacitors need resistors to perform their function, like regular filters?
- & Noun form of "extant"?
- nouble integral problem with sin
- Approximation of Borel sets by a countable collection of majorants
- What are the advantages and disadvantages of an oblique wing over a swing wing
- Funding government by only printing more money
- If a stock doesn't pay dividends, then why is the stock worth anything?
- A short, brutal riddle
- Did a Kamikaze ever impact the hull of a ship, as opposed to the deck?
- Is it modest to compliment a woman on her modesty?
- Why does Scala allow nested data structures like List or Array
- Password strength of random words and brainflaver
- Proving that all integers are even or odd
- My 4.5 yrs old son has no dominant hand
- How to ask dumb questions
- Could a communication satellite that fails to reach GEO be repurposed?
- Did Arnold Rimmer kill the Red

more hot questions

a question feed



more (9)

site design / logo © 2015 Stack Exchange Inc; user contributions licensed under <u>cc by-sa 3.0</u> with <u>attribution required</u>