Graph database schema design - Is this suitable for neo4j?

3

1

Scenario: a simple address book where a user can create his own contacts and organize them by adding them in groups. A contact may have multiple addresses.

I have created the following diagram: ![schema-design][1]

I want to query all the contacts who are placed in group x and live in country y.

Is this schema design good enough for those purposes (I want to use the neo4j database)?

database neo4j database-schema graph-databases schema-design

share improve this question

edited Apr 2 at 21:47

asked Apr 1 at 7:50
Moody
122 ● 9

add a comment

asked 4 months ago
viewed 77 times
active 4 months ago

Blog

Why Stack Overflow is a Good Workplace for Women

2 Answers

active oldest votes

3

It looks like the notion of `country` should be a first class citizen in your graph since your query depends on it. Graph model design typically gets influenced a lot by your query patterns.

So I suggest to have a node labeled `Country` for each country and connect the `Address` node with `:LOCATED_IN` relationships to the country. (consequently drop the country property from the address nodes).

With that change your query is as easy as:

```
MATCH (:Group{name:'family'})<-[:placed_in_group]-(contact)-[:lives-at]->()-[:LOCATE
RETURN contact
```

share improve this answer

answered Apr 1 at 8:14
Stefan Armbruster
23.1k ● 2 ● 28 ● 51

Thank you! I appreciate it. – Moody Apr 1 at 8:30

By the way, Stefan, do you by any chance have any comparison of statistics concerning how much do dense nodes affect performance in cases like this? Say I have ~20K nodes with the `Prop` property (maybe indexed on it as well) and I'm querying the DB just to find all the nodes which have the `Prop` and then filter on the concrete `Prop` value (the amount of possible Prop values is rather small: I can enumerate them all (say 1000 possible values)). Would it be a better design if I create a label `Prop` and create 20K relationships to it? What would be a better choice performance-wise? – cdshines Apr 2 at 12:33

you need to be aware that you might run into locking issues if you connect concurrently a lot of addresses to the same country. Creating a relationship means locking its start and end node as well. However you can easily work around that by having couple of "docking nodes" around the country. The address then selects one docking node e.g. by consistent hashing on the thread id - this prevents locking issues. On the query side I don't have measurements available. – Stefan Armbruster Apr 2 at 12:49

add a comment

Related

101 Has anyone used Graph-based Databases (http://neo4j.org/)?

4 Good graph database for finding intersections (Neo4j? Pegasus? Allegro? …)

12 Storing multiple graphs in Neo4J

2 Neo4j graph database helping to store several time based data values, result of simulation

9 Graph Database Design Methodologies

1 Building a many-to-many db schema using only an unpredictable number of foreign keys

1 which role-permission database design?

0 Better ways to model this simple schema design (MongoDB)?

0 Schema design simple address book Hbase/Cassandra

1 Cassandra (CQL) schema/tables look the same as RDBMS for my scenario

Hot Network Questions

2

One option is to have another node from address for country, as pointed out by stefan Armbruster. If you do not want to change the data structure, just add an index to the field "country" of Address. Then you can have a query

```
MATCH (:Group{name:'family'})<-[:placed_in_group]-(contact)-[:lives-at]->
(:Address{country:'US'})
```

share improve this answer

answered Apr 1 at 8:28

gaurav.singharoy
662 ● 6 ● 18

Thank you. I appreciate the tip! – Moody Apr 1 at 8:30

add a comment

Your Answer

Sign up or log in

Sign up using Google

Sign up using Facebook

Sign up using Stack Exchange

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 database neo4j database-schema graph-databases schema-design or ask your own question.

What is aliasing and what causes it?
Why does Scala allow nested data structures like List or Array
Is it possible to automate tests for in progress sprint
What are the advantages and disadvantages of an oblique wing over a swing wing
How can I grant a non-Super User group permission to force check-in of articles?
Would allowing shorter passwords sometimes be more secure?
Sign that word!
What are some techniques to avoid touching a wet tent wall?
Capture output without redirection and leave it on terminal too
How to deal with people looking at their mobile phone during my presentation?
Why are carpenter's pencils flat?
Using siunitx, \ohm results in an italic Omega
Noun form of "extant"?
Which Star Trek character appeared on screen with the most different ranks?

more hot questions

question feed