

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 x

Optimizing neo4j lucene index

asked 4 months ago

viewed 52 times

active 4 months ago

I set up the fulltext autoindex in my Neo4j db.

0

To be clear, in this case, Neo4j is always using Lucene to do the index. I have around 20 million nodes at the moment and that might soon increase to over 40 million nodes.

▼

For most of the queries performance is fine, almost instant, but sometimes queries like

★

```
"*term"
```

take up to 20 seconds to return.

Could you share some tips about optimizing Neo4j and Lucene to perform faster full text searches? Maybe I should modify some caching properties?

The basic config is well explained in the docs but any sort of well written guides on how to configure and/or alter Lucene's behaviour inside Neo4j are hard to find.

optimization indexing lucene neo4j full-text-search

share improve this question

edited Apr 14 at 15:34

asked Apr 14 at 15:25

Tomazz
1

Perhaps you can share your datamodel and query/code ? – [Michael Hunger](#) Apr 14 at 21:46

add a comment

1 Answer

active

oldest

votes

▲

1

I think your main problem is that you're using a leading wildcard there. [See other answers about lucene performance in general on leading wildcards](#).

▼

If you're only looking for a simple term, you may want to extract terms from documents and link them to separate nodes by that term, so that you can exploit graph connections to get to documents containing a term.

No matter how you build your index, this query `"term"` has to look at just about every possible substring of every string in your database, which is going to take a long time.

You may want to look into tokenizing your documents and extracting those key terms, so that you in the end have something like this:

```
(d:Document)-[:contains]->(t:Term { term: "foo" });
```

Then, when you want to know which documents have "foo" in them, you don't do Lucene anymore, but this:

```
MATCH (t:Term {term: "foo"})<-[:contains]-(d:Document)
RETURN d;
```

I expect this would be much, much, much faster but would require you to do that term extraction on the front end. It will also work mostly for simple terms, and not queries like `[foo?o?o?bar]`

share improve this answer

answered Apr 14 at 16:25

FrobberOfBits
8 227 17 100

Blog

Why Stack Overflow is a Good Workplace for Women

Linked

4 Understanding Lucene leading wildcard performance

Related







- 13 Should an index be optimised after incremental indexes in Lucene?
- 0 Neo4j indexing with Lucene and query with SOLR
- 3 Storing multiple values in single property of Neo4j Lucene index
- 5 Neo4j indexing (with Lucene) - good way to organize node "types"?
- 0 Integer index behavior in neo4j + lucene
- 1 Neo4j/Lucene - How to create a unique index?
- 1 Neo4j - Querying with Lucene
- 0 Neo4j - Single Lucene index or one per user?
- 0 Mixed fulltext and exact matching on neo4j index
- 0 neo4j paging with lucene index OR how to skip index hits

Hot Network Questions

- Compact way of writing $(a + b == c \text{ or } a + c == b \text{ or } b + c == a)$
- Need to convert date format in apex
- Seven overlapping circles
- As a professional photographer, how can one handle a wedding photographer being rude or demanding in a way that will compromise overall image quality?
- What is aliasing and what causes it?
- Does adult neurogenesis occur only in the olfactory bulb and the hippocampus?




Your Answer

B *I*          

-  I added more water to my ciabatta dough by accident
 -  If static methods can't be overridden, how its working here (For Java)?
 -  Why are carpenter's pencils flat?
 -  Draw horizontal lines like open and close brackets
 -  Is it possible to automate tests for in progress sprint
 -  Suggested alternatives for that horrible new noun 'nice-to-have'?
- [more hot questions](#)

Sign up or [log in](#)

Post as a guest

-  Sign up using Google
-  Sign up using Facebook
-  Sign up using Stack Exchange

Name

Email

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 [optimization](#) [indexing](#) [lucene](#) [neo4j](#) [full-text-search](#) or [ask your own question](#).

 [question feed](#)

[tour](#) [help](#) [blog](#) [chat](#) [data](#) [legal](#) [privacy policy](#) [work here](#) [advertising info](#) [mobile](#) [contact us](#) [feedback](#)

TECHNOLOGY

Stack Overflow
Server Fault
Super User
Web Applications
Ask Ubuntu
Webmasters
Game Development
TeX - LaTeX

Programmers
Unix & Linux
Ask Different (Apple)
WordPress Development
Geographic Information Systems
Electrical Engineering
Android Enthusiasts
Information Security

Database Administrators
Drupal Answers
SharePoint
User Experience
Mathematica
Salesforce
ExpressionEngine@Answers
more (13)

LIFE / ARTS

Photography
Science Fiction & Fantasy
Graphic Design
Movies & TV
Seasoned Advice (cooking)
Home Improvement
Personal Finance & Money
Academia
more (9)

CULTURE / RECREATION

English Language & Usage
Skeptics
Mi Yodeya (Judaism)
Travel
Christianity
Arqade (gaming)
Bicycles
Role-playing Games
more (21)

SCIENCE

Mathematics
Cross Validated (stats)
Theoretical Computer Science
Physics
MathOverflow
Chemistry
Biology
more (5)

OTHER

Stack Apps
Meta Stack Exchange
Area 51
Stack Overflow Careers