stackoverfloooooooow

Questions | Tags | Users | Badges | Unanswered | Ask Question

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]   ✕

What's the optimum configuration for sub second query latency in Neo4j?

▲
1
▼
★

I have imported nodes.tsv (350MB, 18M rows, 3 cols) and rels.tsv (5GB, 150M rows, 2 cols) using the batch-importer script. These are my batch.properties file entries

- • neostore.nodestore.db.mapped_memory=250M • neostore.relationshipstore.db.mapped_memory=1000M • neostore.relationshipgroupstore.db.mapped_memory=10M • neostore.propertystore.db.mapped_memory=500M • neostore.propertystore.db.strings.mapped_memory=500M • neostore.propertystore.db.arrays.mapped_memory=215M • dump_configuration=true

I have turned on auto update and auto indexing in ne04j.properties as follows

- • allow_store_upgrade=true • node_auto_indexing=true • node_keys_indexable=name,title • relationship_auto_indexing=true • relationship_keys_indexable=sent_date,has_read

I'm using neo4j 2.2 version on 64 bit windows server that has 1 TB SSD and 256GB ram. What's the configuration for batch importer and neo4j server that I should use for maximum query and data loading peformance? This query for ex: is timing out in the browser

```
MATCH ()-[r:BELONGS_TO]->() RETURN r
```

neo4j

share improve this question

asked Apr 7 at 2:13
user330612
320 ● 1 ● 6 ● 21

add a comment

---

2 Answers

active   oldest   **votes**

▲
1
▼

With millions of nodes that query might be slow no matter what you do, though with the amount of memory you have available maybe it wouldn't be a big deal. This is a good guide for calculating memory settings:

http://neo4j.com/developer/guide-performance-tuning/

✔ While you're playing, I would set the query timeout on the server so that your queries can't jam up the server and force you to need to restart it:

http://neo4j.com/docs/stable/server-configuration.html

You might try starting with `LIMIT` clauses on your queries so that you can get an idea for how the performance degrades as the LIMIT increases.

If you can possibly find a way to limit your query based on node selects that would also be helpful, especially if you can do it by a label or by a label/property combination (which you can index).

Lastly, I would try using `EXPLAIN` in the web console to get an idea for how your queries will be executed:

http://neo4j.com/docs/2.2.0/how-do-i-profile-a-query.html

Also you can use `PROFILE`, though that will run the query, so you'll need to be a bit more careful there. You can probably use the `LIMIT` here too to play and see how things work

share improve this answer
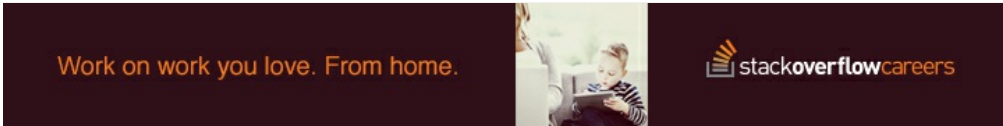
answered Apr 7 at 13:39

Related

1 Neo4j Configuration
4 All queries are slow with neo4j
2 Neo4j Browser – query lost on refresh
0 Neo4j cache configurations
1 Process Crashes when importing 5 million rows to Neo4J
2 Neo4j ServerPlugin Configuration
1 Sub-graph Id a good candidate for a label in neo4j
2 Parametrized like query in neo4j
0 Neo4j crashes all the time randomly
0 query on Neo4j in java applications

Hot Network Questions

Funding government by only printing more money
Make subscript size smaller (always)
Prove the theorem on analytic geometry in the picture.
Carrying a handgun in other countries with a US concealed carry permit
Capture output without redirection and leave it on terminal too
Name for a device purposefully put together from faulty parts
Aerofoil that gives reasonably good lift for both flow directions: forward and backward
Construct a polynomial with given roots
Dullness vs. going overboard: Should I be calling people 'enfants terribles' in an academic paper?

Thanks Brian for the info. My next step is to profile the queries. What's strange is that my overall database folder size is 11.6GB after insertion and index creation. So all of this can easily reside in the RAM I have. And so I'm surprised that queries are timing out for a dB that can reside entirely in memory. I want to know if there are any props I should increase in batch.properties or neo4j – user330612 Apr 7 at 15:02

I think the main problem is that queries need to anchor on nodes, and if you're just querying on `()` then it essentially has to search the entire database and check every relationship on every node for `BELONGS_TO` relationships – Brian Underwood Apr 7 at 16:11

add a comment

▲

2

▼

If you have that much RAM:

Your memory mapping config is wrong for 2.2

use only this setting: `dbms.pagecache.memory=20G`

and then provide neo4j with 24G heap in neo4j-wrapper.conf

Use Neo4j enterprise which scales much better.

Disable the auto-indexes they are not used for what you're doing.

Your query doesn't make sense for any use-case:

```
MATCH ()-[r:BELONGS_TO]->() RETURN r
```

Sub-second graph queries always start at a set of concrete start-points (retrieved with an index lookup) and then traverse out from those starting points.

Global scan queries like your's will just pull all data in memory and inefficiently work over it.

Esp. if you return so much data you can't assume sub-second performance. The data volume alone will kill it.

So figure out a label + property-values that you want to start from and then write the query that traverses out from those start points.

If you want to have sub-second on something you're doing you have to go down to the Java API and aggregate there, e.g. with a server extension:

```
int counter=0;
for (Relationship r : GlobalGraphOperations.at(db)) {
   if (r.hasType(Types.BELONGS_TO)) counter++;
}
return counter;
```

share improve this answer

answered Apr 7 at 22:22

Michael Hunger
25.2k ● 2 ● 21 ● 41

Thanks Michael. I can change the setting of page cage memory in neo4j.propeties files. Where is the . conf file located that I can modify the heap size of vm? Couldn't find any . conf file in my db folder – user330612 Apr 7 at 22:35

it's in the conf folder, if you are on windows check the "Options" button of the starter-app. – Michael Hunger Apr 9 at 21:01

add a comment

Your Answer

B *I* ⊘ 66 {} 🖼 ≣ ≣ ≣ ≣ ↶ ↷

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.

📶 question feed

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