

Setup of mongo and elasticsearch

Clément Levallois

2017-04-10

Table of Contents

| | |
|--|---|
| 1. System | 1 |
| 2. Why MongoDB and Elasticsearch? | 1 |
| 3. Installing MongoDB | 1 |
| 3. Install Elasticsearch | 3 |
| 4. Config Elasticsearch | 4 |
| 5. Install the mongo to elasticsearch connection | 4 |
| a. elastic2-doc-manager | 4 |
| b. run Mongo as a replicaset | 5 |
| 6. Install kibana | 5 |
| 7. Install X-pack | 6 |
| 8. Disable the security component of X-Pack | 6 |
| 9. Install the Mongo-connector for ElasticSearch | 6 |
| 10. Start elasticsearch and Kibana | 7 |

last modified: 2017-10-11

1. System

- I use Debian, version 8.7 ([why?](#))
- Vi is used as a text editor in the following
- **MongoDB 3.4 and Elasticsearch 5.x**

2. Why MongoDB and Elasticsearch?

- MongoDB is a database which stores data without the need for a pre-established model ("strict description") of this data. In practice: I can save something into MongoDB without spending time creating tables and stuff. Just save a JSON doc, that's it.
- MongoDB alone is great, but I will store gigabytes of data, with several text fields and some simple graph logic as well. Elasticsearch is known for managing well the indexes and queries related to these data types.
- A blog post which details how Elasticsearch helped on performances for Mongo: <http://blog.quarkslab.com/mongodb-vs-elasticsearch-the-quest-of-the-holy-performances.html>
- And Elasticsearch can then integrate with [Kibana](#), a way to visualize query results. Awesome!

3. Installing MongoDB

source: <https://docs.mongodb.com/manual/tutorial/install-mongodb-on-debian/>

```
sudo apt-key adv --keyserver hkp://keyserver.ubuntu.com:80 --recv  
0C49F3730359A14518585931BC711F9BA15703C6  
echo "deb http://repo.mongodb.org/apt/debian jessie/mongodb-org/3.4 main" | sudo tee  
/etc/apt/sources.list.d/mongodb-org-3.4.list
```

```
sudo apt-get update
```

```
sudo apt-get install -y mongodb-org
```

disable Transparent Huge Pages as per <https://docs.mongodb.com/manual/tutorial/transparent-huge-pages/>

create a new file:

```
sudo vi /etc/init.d/disable-transparent-hugepages
```

paste this in the text editor:

```
#!/bin/bash
### BEGIN INIT INFO
# Provides:          disable-transparent-hugepages
# Required-Start:    $local_fs
# Required-Stop:
# X-Start-Before:    mongod mongodbmms-automation-agent
# Default-Start:     2 3 4 5
# Default-Stop:      0 1 6
# Short-Description: Disable Linux transparent huge pages
# Description:       Disable Linux transparent huge pages, to improve
#                   database performance.
### END INIT INFO

case $1 in
start)
    if [ -d /sys/kernel/mm/transparent_hugepage ]; then
        thp_path=/sys/kernel/mm/transparent_hugepage
    elif [ -d /sys/kernel/mm/redhat_transparent_hugepage ]; then
        thp_path=/sys/kernel/mm/redhat_transparent_hugepage
    else
        return 0
    fi

    echo 'never' > ${thp_path}/enabled
    echo 'never' > ${thp_path}/defrag

    re='^[0-1]+$'
    if [[ $(cat ${thp_path}/khugepaged/defrag) =~ $re ]]
    then
        # RHEL 7
        echo 0 > ${thp_path}/khugepaged/defrag
    else
        # RHEL 6
        echo 'no' > ${thp_path}/khugepaged/defrag
    fi

    unset re
    unset thp_path
    ;;
esac
```

Make the file executable:

```
sudo chmod 755 /etc/init.d/disable-transparent-hugepages
```

Make the file to be ran on reboot:

```
sudo update-rc.d disable-transparent-hugepages defaults
```

Start Mongo:

```
sudo service mongod start
```

Check that it runs fine:

```
sudo cat /var/log/mongodb/mongod.log
```

→ There should be a line "[initandlisten] waiting for connections on port <port>"

And now stop it, as we will need to run it differently for elasticsearch:

```
sudo service mongod stop
```

3. Install Elasticsearch

source: <https://www.elastic.co/guide/en/elasticsearch/reference/current/deb.html>

Download the public signing key:

```
wget -qO - https://artifacts.elastic.co/GPG-KEY-elasticsearch | sudo apt-key add -
```

Then:

```
sudo apt-get install apt-transport-https
```

```
echo "deb https://artifacts.elastic.co/packages/5.x/apt stable main" | sudo tee -a /etc/apt/sources.list.d/elastic-5.x.list
```

```
sudo apt-get update
```

```
sudo apt-get install elasticsearch
```

```
sudo /bin/systemctl daemon-reload
```

```
sudo /bin/systemctl enable elasticsearch.service
```

4. Config Elasticsearch

```
sudo vi /etc/elasticsearch/elasticsearch.yml
```

→ switch this param to true: bootstrap.memory_lock: true

You then need to make sure the JVM Heap size is no more than half the RAM. First fix a memory param:

```
sudo mkdir /etc/systemd/system/elasticsearch.service.d  
cd /etc/systemd/system/elasticsearch.service.d
```

Add these lines:

```
LimitMEMLOCK=infinity
```

Adjust resource limits:

```
sudo vi /etc/security/limits.conf
```

Add line:

```
elasticsearch - nofile 65536
```

Add a jvm parameter:

```
sudo vi /etc/elasticsearch/jvm.options
```

Add this line:

```
-Djava.io.tmpdir=/var/tmp
```

5. Install the mongo to elasticsearch connection

a. elastic2-doc-manager

This is a doc manager by mongodb labs.

Source: <https://github.com/mongodb-labs/elastic2-doc-manager>

```
sudo apt-get install python-setuptools
sudo easy_install pip
sudo pip install 'elastic2-doc-manager[elastic5]'
sudo pip install 'mongo-connector[elastic5]'
```

b. run Mongo as a replicaset

```
sudo service mongod stop
```

Create the path for your db (if needed)

```
sudo mkdir -p /data/db
```

```
sudo vi /etc/mongod.conf
```

Change dbPath to /data/db

Then:

```
sudo chown -R mongodb:mongodb /data/db
```

Then launch mongo as a replicaset:

```
sudo mongod --port 27017 --dbpath /data/db --replSet rs0 --fork --logpath
/var/log/mongodb.mongod.log
```

6. Install kibana

Kibana is the visualization engine for elastic.

```
sudo wget -q0 - https://artifacts.elastic.co/GPG-KEY-elasticsearch | sudo apt-key add
-
sudo apt-get install kibana
```

Configure Kibana to start automatically at boot:

```
sudo /bin/systemctl daemon-reload
sudo /bin/systemctl enable kibana.service
```

7. Install X-pack

```
https://www.elastic.co/guide/en/x-pack/current/installing-xpack.html
```

- it might need to create an empty file named `/etc/default/elasticsearch`)
- see <https://discuss.elastic.co/t/installing-x-pack-with-nonstandard-conf-dir/76448/3>

INFO

the second command (x-pack install for kibana) takes long minutes, that's normal.

```
cd /usr/share/elasticsearch sudo bin/elasticsearch-plugin install x-pack
```

```
cd /usr/share/kibana sudo bin/kibana-plugin install x-pack
```

8. Disable the security component of X-Pack

This security component is hard to configure, and we don't need it if we run elasticsearch behind a web server and a reverse proxy, on a single machine.

Add `xpack.security.enabled: false`

to `/etc/elasticsearch/elasticsearch.yml`

and to `/etc/kibana/kibana.yml`

Also in the same `kibana.conf` file, change the default username and passwd to "elastic" and "changeme" **and leave the quotes**

- start Elasticsearch: `sudo /usr/share/elasticsearch/bin elasticsearch`
- start Kibana: `sudo /usr/share/kibana/bin kibana`

9. Install the Mongo-connector for ElasticSearch

Source: <https://blog.jixee.me/how-to-use-mongo-connector-with-elasticsearch/>

```
sudo apt-get install python2.7 python-pip curl
sudo pip install mongo-connector
```

Edit the conf of Mongo to turn on replicaset:

```
sudo vi /etc/mongo.conf
(can also be: sudo vi /etc/mongod.conf)
```


Uncomment "replication", add two lines:

```
replication:  
  replSetName: rs0  
  oplogSizeMB: 100
```

```
sudo mongo-connector -m localhost:27017 -t localhost:9200 -d elastic2_doc_manager -n  
database1.collection1,database1.collection2
```

10. Start elasticsearch and Kibana

```
sudo service elasticsearch start  
sudo systemctl start kibana.service
```

You can check that the connection is made here, your Mongo collections should be listed on this page:

```
http://localhost:9200/_cat/indices?v
```

```
== The end  
//ST: The end  
//ST: !
```

Find references for this lesson, and other lessons, [here](#).



This course is made by Clement Levallois.

Discover my other tutorials and courses in data / tech for business: <http://www.clementvallois.net>

Or get in touch via Twitter: [@seinecle](#)