bin/zookeeper-server-start.sh config/zookeeper.properties

bin/kafka-server-start.sh config/server.properties

bin/kafka-topics.sh --list --zookeeper localhost:2181

bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic mytopic --from-beginning

bin/kafka-topics.sh --zookeeper localhost:2181 --describe --topic mytopic

bin/kafka-topics.sh --zookeeper localhost:2181 --delete --topic mytopic

bin/kafka-topics.sh --zookeeper localhost:2181 --alter --topic mytopic --config retention.ms=1000

ps : for influxdb :

bind port 9088

http service port 9086

service : ./influxd --config influxdb.conf

shell : ./influx --port 9086

source ~/.bashrc

for f in ./\*.sh; do source $f; done

source anaconda3/anaconda.sh

source hadoop/hadoop.sh

source spark/spark.sh

comparison :

https://elib.uni-stuttgart.de/bitstream/11682/9793/1/Diplomarbeit\_Comparison\_of\_Time\_Series\_Databases.pdf

https://db-engines.com/en/system/Cassandra%3BKairosDB

Validation :

https://www.scnsoft.com/blog/guide-to-data-quality-management

https://www.cdc.gov/ncbddd/hearingloss/documents/dataqualityworksheet.pdf

/work/weiy/local/anaconda3/bin:/usr/lib64/qt-3.3/bin:/usr/soft/bin:/usr/local/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/work/weiy/.local/bin:/work/weiy/bin:/work/weiy/local/mongodb-linux-x86\_64-rhel70-4.2.3/bin

local termial :

./mongod --config mongodb.conf

./mongo --port 28018

service :

./mongod --port 28018 --dbpath /work/weiy/local/mongodb-linux-x86\_64-rhel70-4.2.3/bin/data/db

## Push some test data using:

##

## curl -H 'X-Warp10-Token: YqeJwThMjxxVCOajB8.b1rmXwIyZh3cEvrapXTC4AlICsxeVNXXKI2MaFvLjAWI917YMAyCKo\_vbPbe6fNW7dHLDQDMJRNhhpUC9AieaGwz.deOeqFUPn7wcbRVBNxYT' http://127.0.0.1:10100/api/v0/update --data-binary '// test{} 42'

##

## And read it back using:

##

## curl 'http://127.0.0.1:10100/api/v0/fetch?token=jMlmjLdrlWWeZMcopcCp\_yskYv4cKWkl1JhkLrU5JXH2LleQchMuqoewJorhmloQ.rvVKU0M\_sfowHHyKEGnFIEHeaDHEsxNAf8CrCUHqVM60dUL0ZXOZko4l8a3Y9u7JF4wU0oHuurFZ\_r9fk2n\_BKVqJNpTlNblYx\_VFt3jksbN3LwAClQkk&selector=~.\*\{\}&now=now&timespan=-1'

##

## You can submit WarpScript for execution via:

##

## curl http://127.0.0.1:8080/api/v0/exec --data-binary @path/to/WarpScriptFile

#-------port 10100

mongodb://localhost:28018/admin

db.setProfilingLevel(0); db.system.profile.drop(); db.setProfilingLevel(2);

m3 :

./bin/m3dbnode -f ./src/dbnode/config/m3dbnode-local-etcd.yml

conda upgrade --all -y

data validation : Cerberus (based on PyPi?)

pip install cerberus

Successfully installed cerberus-1.3.2

data visualisation :

pip install visdom

Installing collected packages: jsonpointer, jsonpatch, torchfile, websocket-client, visdom

Successfully installed jsonpatch-1.25 jsonpointer-2.0 torchfile-0.1.0 visdom-0.1.8.9 websocket-client-0.57.0

python -m visdom.server or visdom

pip install jaeger-client

opentracing-instrumentation-3.2.1 tornado-5.1.1

pip install opentracing\_instrumentation

jaeger-client-4.3.0 opentracing-2.3.0 threadloop-1.0.2 thrift-0.13.0

change jupyter-notebook port :

jupyter-notebook --port 9888

jaeger-1.17.0-linux-amd64

./jaeger-all-in-one --collector.zipkin.http-port=9411

# to build a local m3dbnode process

make m3dbnode (note that we currently require at least Go 1.10 or higher)

# run it with the sample configuration

./bin/m3dbnode -f ./src/dbnode/config/m3dbnode-local-etcd.yml

curl -X POST http://localhost:7201/api/v1/database/create -d '{

"type": "local",

"namespaceName": "default",

"retentionTime": "12h"

}'

Lundi :

- installation du logiciel : Hadoop + Spark

- Comprendre la notion de pyspark

- Configuration findspark

- Tps de pyspark

Mardi :

- Notion de hadoop

- Tps ML en pyspark

- Notion de pipline

Mercredi :

- Installation du logiciel : KairosDB + appach Kafka

- Notion de Zookeeper, topic, ...

-

Difficultés :

configuration d’environnement

variables d’environnement dans spyder

pas de droit d’exécuter certains commandes

Dépendance :

kafka\_2.12-2.4.0.jar

configuration kairosdb :

créer répertoire var dans kairosdb pour fichier PID

kairosdb.telnetserver.port = 4343

HTTP port : kairosdb.jetty.port=9090

probleme de code 504

influxdb

InfluxDB shell version: 1.7.10

bind-address = "127.0.0.1:9088"

Mongodb :

logpath: /work/weiy/local/var/mongodb/log/mongod.log

dbPath: /work/weiy/local/var/mongodb/lib/mongo

pidFilePath: /work/weiy/local/var/run/mongodb/mongod.pid

config : mongodb.conf

port 28018

problème importer librairies libcrypto.so.100

change version of mongodb

M3 :

Installation go

tutorial :

https://github.com/m3db/m3/blob/master/DEVELOPER.md

Connecteur :

python -m pip install pymongo

Connecter Wrap10 :

pip install warp10-jupyter

warp10-jupyter-1.0

pip install py4j

Installing collected packages: py4j

Successfully installed py4j-0.10.9

visual tools

mongodb compass 1.20.5

rpm2cpio mongodb-compass-1.20.5.x86\_64.rpm | cpio -idv