

TP NOSQL

- 1) Pour chaque modèle de données 5 instances de machines seront déployées sur des VMs Centos ou sur Docker.

Lien de téléchargement d'image centos7 ou centos8 pour Virtual box ou VMware

<https://www.linuxvmimages.com/images/centos-7/>

<https://www.linuxvmimages.com/images/centos-8/>

Pour ceux qui aimerait faire de la conteneurisation vous pouvez vous aider des liens suivant :

https://hub.docker.com/_/redis

https://hub.docker.com/_/mongo

<https://neo4j.com/developer/docker/>

https://hub.docker.com/_/cassandra

<https://medium.com/@michaeljpr/five-minute-guide-getting-started-with-cassandra-on-docker-4ef69c710d84>

<https://ralph.blog.imixs.com/2020/06/22/setup-a-public-cassandra-cluster-with-docker/>

<https://neo4j.com/developer/docker-run-neo4j/>

<https://medium.com/@ManagedKube/deploy-a-mongodb-cluster-in-steps-9-using-docker-49205e231319>

- 2) Pour chaque modèle quand c'est possible implémenter :

- **Le sharing**
- **Le sharding**
- **La tolérance aux pannes**
- **La Consistance**
- **La disponibilité**

- 3) Pour chaque modèle s'inspirer des liens pour installer les solutions en mode cluster
NB : Ces liens ne sont pas obligatoires vous pouvez effectuer vos propres recherches

- **Modèle de données clé-valeur (de préférence Redis)**

<https://redis.io/>

<http://tgrall.github.io/blog/2019/09/05/multi-nodes-redis-cluster-with-docker/>

<https://www.tecmint.com/setup-redis-cluster-in-centos-8/>

<https://www.digitalocean.com/community/tutorials/how-to-configure-a-redis-cluster-on-centos-7>

<https://linuxize.com/post/how-to-install-and-configure-redis-on-centos-7/>

<https://blog.opstree.com/2019/10/29/redis-cluster-setup-sharding-and-failover-testing/>
<https://redis.io/topics/cluster-tutorial>

■ **Modèle de données document (de préférence MongoDB ou Elasticsearch)**

<https://www.mongodb.com/>

<https://blog.jscribler.com/how-we-achieved-mongodb-replication-on-docker/>

<https://www.howtoforge.com/tutorial/install-a-mongodb-sharded-cluster-on-centos-7/>

<https://medium.com/@thepradeep001/mongodb-advance-operations-5755693fd76e>

■ **Modèle de données graphe (de préférence Neo4j ou tigergraph)**

<https://medium.com/@prabhin.mp/how-to-install-neo4j-on-centos-d5c1be242471>

<https://neo4j.com/docs/operations-manual/current/installation/linux/rpm/>

<https://neo4j.com/docs/operations-manual/current/clustering/>

<https://neo4j.com/docs/operations-manual/current/tools/cypher-shell/>

■ **Modèle de données colonne (de préférence Cassandra)**

<https://cassandra.apache.org/>

<https://www.digitalocean.com/community/tutorials/how-to-run-a-multi-node-cluster-database-with-cassandra-on-ubuntu-14-04>

<https://intercax.atlassian.net/wiki/spaces/SYN32/pages/283213841/Apache%2BCassandra%2BSingle%2BMulti-Node%2BCluster%2BSetup%2BTesting%2BInstructions%2Bfor%2BRHEL%2BCentOS7>

<https://blog.knoldus.com/a-simple-walk-through-to-set-up-a-local-cassandra-multi-node-cluster/>
<https://knowledge.broadcom.com/external/article/9929/howto-configure-a-multinode-cluster-with.html>

<https://www.linode.com/docs/guides/set-up-a-cassandra-node-cluster-on-ubuntu-and-centos/>

<https://www.howtoforge.com/tutorial/how-to-install-cassandra-database-on-centos-7/>

- 4) Choisir un jeu de données de votre choix pour effectuer les opérations ci-dessous
NB : Ces liens ne sont pas obligatoires vous pouvez effectuer vos propres recherches

Large Dataset :

<https://www.stats.govt.nz/large-datasets/csv-files-for-download/>

<https://graphchallenge.mit.edu/data-sets>

<https://lgylm.github.io/big-graph/dataset.html>

<https://www.kaggle.com/datasets>

<https://snap.stanford.edu/data/>

Effectuer des opérations de :

- Création de schéma quand c'est possible
- Chargement de données
- CRUD (create, read, update and delete)
- Faire du requetage sur les différentes collections

NB : Vous pouvez bien choisir une solution NoSQL de votre choix suivant les modèles vous pouvez vous aider de ce lien <https://db-engines.com/en/ranking>