



Cati Sicpa

Ce document est mise à disposition selon les termes de la
[Licence Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

Installation de Apache Cassandra sous CentOS 7

19 avril 2018

1 Objectifs

Cassandra est un système libre de gestion de bases de données distribuées qui peut être utilisé pour stocker des gros volumes de données à travers plusieurs servers. Cassandra prend en charge la réplication dans plusieurs data centers pour garantir la redondance, le basculement et la reprise après un sinistre. Ce document est le premier d'une série dont le but est de mettre en place un cluster Apache Cassandra formé de 2 data centers ayant chacun de 2 racks. Chaque rack regroupe 2 machines virtuelles ayant CentOS 7 pour système d'exploitation. Les caractéristiques de base de ces machines sont : 8 GB de RAM et 32 GB de disque dur. Le cluster à mettre en place est donc constitué de 8 machines virtuelles CentOS 7. La première étape est d'installer Apache Cassandra sur chacune de ces 8 machines virtuelles. Ce document décrit comment installer Cassandra dans une machine CentOS 7.

2 Pré-requis

Avant d'installer Cassandra, il est nécessaire d'avoir déjà installé dans la machine une version correcte de Java et une version récente de Python. Avant de commencer, on peut mettre à jour le système avec la dernière version en exécutant la commande suivante :

```
# sudo yum update -y
```

3 Installation de Java

Cassandra est écrit en Java. ainsi, nous devons avoir Java installé dans le système. Nous avons besoin de la dernière version 64-bit de java 8, soit Oracle Java Standard Edition 8 ou OpenJDK 8. Nous choisissons OpenJDK 8. Nous pouvons l'installer en exécutant la commande suivante :

```
# sudo yum install java -y
```

Après l'installation de Java, nous pouvons vérifier la version avec la commande suivante :

```
# sudo java -version
```

4 Installation de Python

Nous avons besoin de Python pour le fonctionnement du shell cqlsh, l'interface de ligne de commande pour la base de données Cassandra. Plus spécifiquement, nous devons installer la dernière version de Python 2.7. Notons que Python 2.7.5 est déjà installé dans le système CentOS 7 et plusieurs applications du système en dépendent. Sa désinstallation endommagerait le système. Nous allons installer la dernière version Python 2.7.14 sans désinstaller l'ancienne version. Pour cela nous pouvons suivre les étapes suivantes.

4.1 Installer GCC

Pour commencer, il faut vérifier que le package GCC est installé dans notre système. S'il ne l'est pas, on peut utiliser la commande suivante pour l'installer :

```
# yum install gcc openssl-devel bzip2-devel
```

4.2 Télécharger Python 2.7

Utiliser la commande ci-dessous pour télécharger Python à partir du site officiel. On peut y remplacer la version spécifiée par la toute dernière version en cours.

```
# cd /usr/src
```

```
# yum install wget
```

```
# wget https://www.python.org/ftp/python/2.7.14/Python-2.7.14.tgz
```

4.3 Extraction de l'archive et compilation

Utiliser l'ensemble des commandes ci-dessous pour extraire le code source de Python et le compiler. Nous utilisons "altinstall" pour éviter de remplacer le fichier binaire python par défaut "/usr/bin/python"

```
# tar xzf Python-2.7.14.tgz
```

```
# cd Python-2.7.14
```

```
# ./configure --enable-optimizations
```

```
# make altinstall
```

4.4 Vérification de la version de Python installée

Vérifier la version de Python installée en utilisant la commande ci-dessous qui indique aussi le chemin vers le fichier binaire de la nouvelle version de Python.

```
# /usr/local/bin/python2.7 -V
```

5 Installation de Cassandra

Ajouter l'entrepot Apache de Cassandra dans le fichier `"/etc/yum.repos.d/cassandra.repo"`, par exemple la dernière version 3.11. On peut le faire en exécutant la commande suivante :

```
# vi /etc/yum.repos.d/cassandra.repo
```

et insérer dans ce fichier les lignes suivantes :

```
[cassandra]
name=Apache Cassandra
baseurl=https://www.apache.org/dist/cassandra/redhat/311x/
gpgcheck=1
repo_gpgcheck=1
gpgkey=https://www.apache.org/dist/cassandra/KEYS
```

Enregistrer et fermer lorsque cette insertion est terminée. Puis, effectuer la mise à jour du repertoire avec la commande suivante :

```
# sudo yum update -y
```

Installer Apache Cassandra à l'aide le commande suivante tout en acceptant les prompts d'import des clés gpg :

```
# sudo yum install cassandra
```

Mettre en marche Cassandra avec la commande suivante :

```
# service cassandra start
```

Exécuter la commande suivante pour lancer automatiquement Apache Cassandra au démarrage :

```
# chkconfig cassandra on
```

Vérifier le bon fonctionnement de Cassandra avec la commande suivante :

```
# sudo nodetool status
```

Se connecter à Cassandra à l'aide du shell `"cqlsh"` en exécutant la commande ci-dessous :

```
# cqlsh
```

Sortir de ce shell en exécutant la commande suivante :

```
# cqlsh> exit;
```

Stopper Cassandra avec la commande ci-dessous :

```
# sudo service cassandra stop
```

6 Conclusion

Nous avons suivi les instructions présentées dans ce document pour installer Apache Cassandra independamment dans chacun des 8 machines du cluster que nous souhaitons mettre en place. Le prochain document décrira les étapes de configuration de ce cluster.

Références