



INRAE



Cati Sicpa

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

Installation de Apache Solr sous CentOS 7

29 août 2018

1 Objectifs

Apache Solr (lire solar) est un moteur de recherche distribué, hautement disponible et résistant aux défaillances qui est construit sur Apache Lucene. Il offre la possibilité de consulter facilement, rapidement et de différentes manières des informations indexées à l'aide des mots clés ou des expressions. Apache Solr est ainsi une base de données NoSQL orientée recherche. L'indexation et la recherche sont les deux opérations primaires effectuées par Solr. L'indexation est le processus d'insérer les données dans Solr afin de les rendre consultable. Mais avant d'indexer les données, il faut installer et configurer convenablement Solr sur un serveur ou sur un ensemble de plusieurs serveurs distincts si l'on souhaite manipuler une grande quantité de données. Dans ce document, nous montrons comment installer Apache Solr sur un seul serveur. Dans un prochain document, nous expliquerons la configuration d'un cluster de serveurs Solr.

2 Pré-requis

Avant d'installer Solr, il est nécessaire d'avoir déjà installé dans la machine une version correcte de Java. 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

Apache Solr 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 Ouverture des ports TCP

Si le pare-feu est activé dans les machines, il est nécessaire d'ouvrir et de persister quelques ports TCP spécifiques. On peut le faire avec les commandes ci-dessous :

```
# firewall-cmd --permanent --add-port=8983/tcp
# firewall-cmd --permanent --add-port=7574/tcp
# firewall-cmd --permanent --add-port=8984/tcp
# firewall-cmd --permanent --add-port=8985/tcp
# firewall-cmd --permanent --add-port=9983/tcp
# firewall-cmd --permanent --add-port=9984/tcp
# firewall-cmd --permanent --add-port=9985/tcp
# firewall-cmd --permanent --add-port=18983/tcp
# firewall-cmd --permanent --add-port=18984/tcp
# firewall-cmd --permanent --add-port=18985/tcp
# firewall-cmd --reload
```

On peut également vérifier les ports ouverts avec la commande suivante :

```
# firewall-cmd --list-ports
```

5 Installation de Solr

Apache Solr peut être installé dans n'importe quel système d'exploitation. Pour l'installer dans une machine, il faut télécharger et décompresser le fichier `solr-7.4.0.tgz` dans un répertoire quelconque de la machine. Nous pouvons par exemple exécuter les commandes suivantes :

```
# yum install wget
#wget http://wwwftp.ciril.fr/pub/apache/lucene/solr/7.4.0/solr-7.4.0.tgz
#tar -zxf solr-7.4.0.tgz
#cp -r solr-7.4.0 /opt/solr-7.4.0
```

Nous définissons les variables SOLR_HOME et PATH de façon permanente et globale pour tous les utilisateurs. Pour cela, nous éditons le fichier `/etc/profile` de chaque machine avec la commande suivante :

```
# vi /etc/profile
```

Ensuite, nous insérons les trois lignes ci-dessous juste avant la ligne contenant le texte `export PATH USER LOGNAME MAIL HOSTNAME HISTSIZE HISTCONTROL`

```
export SOLR_HOME=/opt/solr-7.4.0/server/solr
export PATH=$PATH:/opt/solr-7.4.0/bin
export PATH=$PATH:/opt/solr-7.4.0/server/scripts/cloud-scripts
```

A la fin de cette modification, nous enregistrons et fermons ce fichier. Ensuite, nous exécutons la commande ci-dessous pour que ces modifications soient immédiatement prises en compte par le système :

```
# source /etc/profile
```

Sous le système d'exploitation CentOS, Apache Solr a besoin de l'application `lsol` pour fonctionner correctement. On peut installer cette application avec la commande suivante :

```
# yum install lsol -y
```

6 Démarrage et arrêt de Solr

A ce stade, Solr est parfaitement installé et prêt à l'utilisation.

On peut consulter la version de Solr qui est installée avec la commande suivante :

```
# solr -version
```

On démarre Solr avec la commande suivante dans laquelle l'option `-force` est nécessaire si on est connecté sur le serveur en tant que root.

```
# solr start -force
```

Une fois que Solr est en marche, on peut consulter son status avec la commande ci-dessous :

```
# solr status
```

Apache Solr offre une bonne interface web pour l'administration l'indexation et la recherche que l'on peut accéder si possible avec le lien `http://localhost:8983/solr` si on est en local ou le lien `http://<ip_address>:8983/solr` sinon, où `<ip_address>` est l'adresse IP de la machine sur laquelle Solr est installé.

Pour terminer, on arrête Solr avec la commande suivante

```
# solr stop
```

7 Conclusion

Dans ce document, nous avons installé Apache Solr sur un Serveur CentOS. Nous envisageons de mettre en place un cluster Solr constitué de 12 machines qui soit configuré de façon à être hautement disponible et sans point unique de défaillance. Le prochain document décrira les étapes d'une telle configuration.

Références