Tuto1 hdp

Version 1.0

2019-01-31

DATASOFT CONSULTING

Historique des modifications du document

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Version** | **Description** | **Auteur** |
| 2019-01-30 | 1.0 | Tuto1\_hdp | Abdelaziz HANI |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Table des matières

1. INTRODUCTION 4

2. ENVIRONMENT SETUP 4

2.1 DETERMINER L'ADRESSE IP DE VOTRE SANDBOX 4

2.2 MAP SANDBOX IP À VOTRE NOM D'HOTE SOUHAITÉ DANS LE FICHIER HÔTES 4

2.3 ACCÈS TERMINAL 5

2.4 MÉTHODE CLIENT WEB CLIENT 5

2.5 ENVOYER DES DONNEES ENTRE SANDBOX ET MACHINE LOCALE 5

2.6 PAGE D'ACCUEIL 5

2.7 EXPLOREZ AMBARI 6

2.7.1 IDENTIFIANT DE CONNEXION: 8

# INTRODUCTION

Ce tutoriel est destiné aux utilisateurs peu expérimentés dans l’utilisation de la Sandbox.

La Sandbox est un environnement d’apprentissage simple, préconfiguré, qui contient les derniers développements d’Apache Hadoop, en particulier la plate-forme Hortonworks Data Platform (HDP). La Sandbox est livrée dans un environnement virtuel pouvant s'exécuter dans le cloud ou sur votre ordinateur personnel. La Sandbox vous permet d'apprendre et d'explorer le HDP par vous-même.

# ENVIRONMENT SETUP

C’est la section administrative pour commencer à utiliser l’environnement Hortonworks Sandbox.

## DETERMINER L'ADRESSE IP DE VOTRE SANDBOX

Une fois que la Sandbox est installé sur l'hôte de votre environnement, l'adresse IP varie en fonction de votre machine virtuelle (VMware, VirtualBox). Une fois que la Sandbox est en cours d'exécution, il vous indiquera l'adresse IP. Un exemple d'adresse IP typique pour l’environnement pris en charge:

VMWare: adresse IP = 192.168.x.x

VirtualBox: adresse IP = 127.0.0.1

Si vous utilisez VirtualBox ou VMWare, vous pouvez confirmer l’adresse IP en attendant la fin de l’installation et un écran de confirmation vous indiquera l’adresse IP résolue par votre sandbox. Par exemple:


## MAP SANDBOX IP À VOTRE NOM D'HOTE SOUHAITÉ DANS LE FICHIER HÔTES

Mac, Linux et Windows ont tous un fichier hosts. Une fois configuré, ce fichier permet de mapper l'adresse IP du sandbox sur un nom d'hôte plus facile à retenir qu'un numéro.

**Mac users:**

* echo '{IP-Address} sandbox-hdp.hortonworks.com sandbox-hdf.hortonworks.com' | sudo tee -a /private/etc/hosts

**Linux users:**

* echo '{IP-Address} sandbox-hdp.hortonworks.com sandbox-hdf.hortonworks.com' | sudo tee -a /etc/hosts

**Windows users:**

* Exécutez le Bloc-notes en tant qu'administrateur.
* Ouvrir le fichier hosts situé dans: c: \Windows\System32\drivers\etc\hosts.
* Ajoutez {IP-Address} localhost sandbox-hdp.hortonworks.com sandbox-hdf.hortonworks.com

Sauvegarder le fichier.

## ACCÈS TERMINAL

Vous pouvez également vous connecter à l'aide de root, à l'aide du mot de passe hadoop, ce qui vous obligera peut-être à changer le mot de passe.

Si vous vous connectez avec des identifiants autres que root, vous devrez utiliser sudo avant la commande. Par exemple: **sudo ambari-server status**

Ouvrez votre terminal (mac / Linux) ou Git Bash (Windows). Tapez la commande suivante pour accèder à la sandbox via ssh user@hostname -p port. Par exemple:

**ssh root@sandbox-hdp.hortonworks.com -p 2222**


## MÉTHODE CLIENT WEB CLIENT

Le client Web shell est également appelé Shell-in-a-Box. C’est un moyen simple d’émettre des commandes shell sans avoir à installer de logiciel supplémentaire. Il utilise le port 4200, par exemple:

sandbox-hdp.hortonworks.com:4200

## ENVOYER DES DONNEES ENTRE SANDBOX ET MACHINE LOCALE

En utilisant le terminal de votre choix, vous pouvez transférer des fichiers de / vers le bac à sable et la machine locale.

Transférer le fichier de la machine locale vers la sandbox:

scp -P 2222 <local\_directory\_file> root@sandbox-hdp.hortonworks.com:<sandbox\_directory\_file>

 Transférer le fichier du sandbox vers la machine locale:

scp -P 2222 root@sandbox-hdp.hortonworks.com:<sandbox\_directory\_file> <local\_directory\_file>

Avez-vous remarqué la différence entre les deux commandes?

Pour envoyer des données de la machine locale au sandbox, le chemin du répertoire de la machine locale vient avant le répertoire du sandbox. Pour transférer les données du sandbox vers la machine locale, les arguments de la commande sont inversés.

## PAGE D'ACCUEIL

La page de bienvenue de Sandbox est également appelée page de démarrage. Il fonctionne sur le numéro de port: 1080. Pour l'ouvrir, utilisez votre adresse d'hôte et ajoutez le numéro de port. Par exemple:

<http://sandbox-hdp.hortonworks.com:1080/>

Cela ressemble à ceci:

**Launch Dashboard** ouvre deux fenêtres de navigateur - l’interface Ambari et le didacticiel pour débutants. Vous devez vous connecter à Ambari en utilisant le nom d'utilisateur et le mot de passe en fonction des exigences du tutoriel. La plupart des tutoriels utilisent raj\_ops ou maria\_dev. Reportez-vous à Informations d'identification pour la liste des utilisateurs et des mots de passe.

**Advanced HDP Quick Links** fournit un accès rapide aux services Ambari tels que Zeppelin, Atlas, Ranger, Shell-in-a-box et autres.

## EXPLOREZ AMBARI

* Ambari Dashboard fonctionne sur le port : 8080. Par exemple : <http://sandbox-hdp.hortonworks.com:8080>
* Connectez-vous en tant qu'administrateur, voir Réinitialisation du mot de passe de l'administrateur.
* Sélectionnez **manage Ambari**

L'écran suivant sera affiché:

1. **« Dashboard »** vous mènera au tableau de bord Ambari, qui est la principale interface utilisateur pour les opérateurs Hadoop.
2. **« Cluster Management »** vous permet d'accorder des autorisations aux utilisateurs et aux groupes Ambari.
3. **« Administration of Users »** vous permet d’ajouter et de supprimer des utilisateurs et des groupes Ambari.
4. **« Ambari User Views »** répertorie l’ensemble des vues d’utilisateurs Ambari faisant partie du cluster.

Cliquez sur **Go to Dashboard** et vous devriez voir un écran similaire:

Explorer en cliquant sur:

1. **Metrics, Heatmaps and Config History**

Et ensuite:

1. **Background Operations, Alerts, Admin and User Views, Icon (**représentée par une matrice 3 × 3**)** pour vous familiariser avec les ressources Ambari mises à votre disposition.

### **IDENTIFIANT DE CONNEXION:**

|  |  |
| --- | --- |
| **USER** | **MOT DE PASSE** |
| **maria\_dev** | maria\_dev |
| **raj\_ops** | raj\_ops |
| **holger\_gov** | holger\_gov |

* **maria\_dev** - Responsable de la préparation et de la compréhension des données. Elle aime explorer différents composants HDP tels que Hive, Pig, HBase.
* **raj\_ops** - Responsable de la construction de l'infrastructure, des activités de recherche et développement telles que la conception, l'installation, la configuration et l'administration. Il est expert technique dans le domaine de l’administration système pour les systèmes d’exploitation complexes.
* **holger\_gov** - principalement pour la gestion des éléments de données, à la fois le contenu et les métadonnées. Il a un rôle de spécialiste qui intègre des processus, des politiques, des directives et des responsabilités pour l’administration de l’ensemble des données de l’organisation, conformément aux exigences de la politique et / ou de la réglementation.