Tuto1 hdf

Version 1.0

2019-01-31

DATASOFT CONSULTING

Historique des modifications du document

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Version** | **Description** | **Auteur** |
| 2019-01-31 | 1.0 | Tuto1\_hdf | Abdelaziz HANI |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Table des matières

1. INTRODUCTION 4

2. QU'EST-CE QU'UNE SANDBOX ? 4

3. HDF STANDALONE SANDBOX 4

3.1 CONFIGURATION DE L'ENVIRONNEMENT 4

3.2 DETERMINER L'ADAPTATEUR RESEAU DE VOTRE SANDBOX 4

3.3 DETERMINER L'ADRESSE IP DE VOTRE SANDBOX 4

3.4 MAP SANDBOX IP À VOTRE NOM D'HOTE SOUHAITÉ DANS LE FICHIER HÔTES 5

3.5 ACCÈS TERMINAL 5

3.6 METHODE SHELL SECURISEE 6

3.7 METHODE DU CLIENT WEB 6

3.8 ENVOYER DES DONNEES ENTRE LA SANDBOX ET LA MACHINE LOCALE 6

4. PAGE D'ACCUEIL 6

4.1 EXPLOREZ AMBARI 7

4.2 IDENTIFIANT DE CONNEXION: 9

4.3 SANDBOX VERSION 9

# INTRODUCTION

La création d'applications liées à l’IOT est plus rapide et plus simple grâce à la structure open source de données en mouvement appelée Hortonworks DataFlow (HDF). Apprenez à créer des applications IOT dans un environnement de test virtuel garantissant la sécurité de votre environnement informatique personnel. Le format HDF peut être appris via une sandbox HDF. Des didacticiels ont été mis au point et testés sur le bac à sable pour faciliter le démarrage de Big Data et d'IOT.

À la fin de ce didacticiel, vous serez familiarisé avec les outils de données en mouvement de la plate-forme HDF qui peuvent être utilisés pour créer vos applications.

# QU'EST-CE QU'UNE SANDBOX ?

La Sandbox est un environnement d'apprentissage simple, préconfiguré, qui contient les derniers développements des outils liés à Apache Big Data, en particulier ces outils ont été assemblés dans Hortonworks DataFlow (HDF). La Sandbox est livrée dans un environnement virtuel pouvant s'exécuter dans le cloud ou sur votre ordinateur personnel. Le bac à sable vous permet d'apprendre à créer des applications de diffusion en continu. Des tutoriels supplémentaires testés par rapport au dernier environnement sandbox sont fournis dans HDF Tutorial Track.

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

# HDF STANDALONE SANDBOX

L'architecture Sandbox HDF Standalone est fournie avec les outils Big Data suivants:

**Zookeeper, Storm, Kafka, NiFi, NiFi Registry, Schema Registry, Streams Messaging Manager (SMM), and Stream Analytics Manager (SAM).**

## CONFIGURATION DE L'ENVIRONNEMENT

C’est la section administrative pour commencer à utiliser l’environnement Hortonworks Sandbox. Généralement, cela ne sera fait qu'une fois.

## DETERMINER L'ADAPTATEUR RESEAU DE VOTRE SANDBOX

Une fois que la machine virtuelle Sandbox VirtualBox est installée, elle se connecte à un réseau virtuel. Il existe 8 modes de réseau différents, mais le réseau par defaut auquel votre sandbox se connectera est NAT. Nous couvrirons les réseaux pertinents pour nos cas d'utilisation tutoriels: NAT et adaptateur ponté.

## 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'adresses IP typiques pour chaque environnement pris en charge:

VirtualBox: adresse IP = 127.0.0.1

VMWare: adresse IP = 192.168.x.x


## 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
* Ajouter :

{IP-Address}localhost sandbox-hdp.hortonworks.com sandbox-hdf.hortonworks.com

* Enregistrez le fichier

IMPORTANT: Replacez **{IP-Address}** avec [**Sandbox IP Address**](https://fr.hortonworks.com/tutorial/learning-the-ropes-of-the-hdf-sandbox/#determine-ip-address-of-your-sandbox)

## ACCÈS TERMINAL

Reportez-vous à Informations d'identification pour la liste des utilisateurs et des mots de passe. 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 de mot de passe - souvenez-vous du mot de passe que vous avez défini!

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

## METHODE SHELL SECURISEE

Ouvrez votre terminal (mac / Linux) ou Git Bash (Windows). Tapez la commande suivante pour accèder au sandbox via **ssh user@hostname -p port**. par exemple: **ssh root@sandbox-hdf.hortonworks.com -p 2202**


## METHODE DU CLIENT WEB

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-hdf.hortonworks.com:4200

## ENVOYER DES DONNEES ENTRE LA SANDBOX ET LA MACHINE LOCALE

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

* Transférer le fichier de la machine locale vers la sandbox :

scp -P 2202 <local\_directory\_file> root@sandbox-hdf.hortonworks.com:<sandbox\_directory\_file>

* Transférer le fichier du sandbox sur la machine locale :

scp -P 2202 root@sandbox-hdf.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-hdf.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 admin. Reportez-vous à Informations d'identification pour la liste des utilisateurs et des mots de passe.

**Advanced HDF Quick Links** fournit un accès rapide aux services Ambari tels que NiFi, SAM,

Schema Registry, 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**et**Config History**
2. **Hosts, Alerts, Admin**

## IDENTIFIANT DE CONNEXION:

|  |  |
| --- | --- |
| **USER** | **MOT DE PASSE** |
| **raj\_ops** | raj\_ops |

* **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.

## SANDBOX VERSION

Lorsque vous rencontrez un problème, l'une des premières choses à poser est la suivante: "Quelle version de sandbox utilisez-vous"? Pour obtenir cette information:

Login avec [Web-Shell-Client](http://sandbox-hdf.hortonworks.com:4200/) et exécutez: **sandbox-version**. La sortie devrait ressembler à quelque chose comme:

