Creating a TechZone-based StreamSets Demo Environment

Table of Contents

[Introduction 2](#_Toc174544799)

[Provision a TechZone-based StreamSets Data Plane Instance 2](#_Toc174544800)

[Choose your deployment method 3](#_Toc174544801)

[Log in to StreamSets Platform 6](#_Toc174544802)

[1. If you already have an account on StreamSets Platform 6](#_Toc174544803)

[2. If you do not yet have an account on StreamSets Platform 6](#_Toc174544804)

[Deploy a StreamSets Engine 8](#_Toc174544805)

[1. Clone the Hands On Lab Deployment 8](#_Toc174544806)

[2. Set Permissions on the Deployment 10](#_Toc174544807)

[3. Start the Deployment 12](#_Toc174544808)

[4. Generate an Engine Deployment Script 13](#_Toc174544809)

[5. Run the Engine Deployment Script 14](#_Toc174544810)

[6. Confirm the Engine is Healthy 15](#_Toc174544811)

[Run a Test Pipeline 16](#_Toc174544812)

# Introduction

This document describes how to provision and configure a TechZone-based StreamSets Demo Environment.

# Provision a TechZone-based StreamSets Data Plane Instance

Provision an instance of the Data Plane for StreamSets VM in TechZone at this link: <https://techzone.ibm.com/collection/streamsets/environments>

A screenshot of a phone

Description automatically generated

Once your TechZone reservation is provisioned, you should see connection information like this (with your own unique port numbers)

A close-up of a website

Description automatically generated

# Choose your deployment method

In the steps that follow, you will be asked to copy a block of text (a StreamSets Engine deployment command) from your Web Browser and to paste that text into a terminal session connected to a TechZone VM. In order for that copy/paste to work, the Browser and the terminal session must both be running either on your local machine (preferred), or, if you are unable to run a terminal session from your desktop machine, from the UI of the TechZone VM.

Here are details of the two deployment methods:

* If using your local machine (preferred), open a terminal session and ssh to the TechZone VM using the example ssh command shown in the screenshot above. Use a browser running on your local machine. Skip the rest of this step and proceed to the next step.
* If running the terminal session and the browser from the TechZone VM's UI, click the Console button in your reservation's detail window:

A screenshot of a computer

Description automatically generated

Then click the Open in a new window button:

A red and black rectangular with a power button

Description automatically generated

Log in as the admin user using the password IBMDem0s

A screenshot of a login box

Description automatically generated

Click the Activities button on the top left, and then the terminal icon below:

A screenshot of a computer

Description automatically generated

→ That will open a terminal session logged in as the admin user:

A black screen with white text

Description automatically generated

Similarly, to open Firefox, click the Activities button on the top left, and then the Firefox icon below:

A screenshot of a computer

Description automatically generated

# Log in to StreamSets Platform

## If you already have an account on StreamSets Platform

If you already have an account with StreamSets Platform, you can log into your account at:

<https://cloud.login.streamsets.com/login>

If prompted, select the IBM StreamSets Hands On Lab Organization.

***→ Once you connect, bookmark the StreamSets Platform home page!***

and then you can skip ahead to the next section “Deploy a StreamSets Engine”.

## If you do not yet have an account on StreamSets Platform

If you do not yet have an account for the StreamSets Platform, when you provision the TechZone VM, you will be added as a user to StreamSets Platform, and you should receive an email like this (please check your spam folder if you do not see the invite):

A screenshot of a computer

Description automatically generated

→ Click the Join button

Your browser should open to this dialog:

A screenshot of a login box

Description automatically generated

***→ Don't worry about the "Try for 30 Days" message; your account will not expire.***

Enter the requested information and you should be connected to the StreamSets Platform home page:

A screenshot of a web page

Description automatically generated

***→ Bookmark the StreamSets Platform home page!***

# Deploy a StreamSets Engine

## Clone the Hands On Lab Deployment

* Navigate to Setup > Deployments and clone the Hands On Lab Deployment:

A screenshot of a computer

Description automatically generated

* Edit the Deployment Name using the naming convention <first\_name>\_<last\_name>\_deployment. For example, I'll name mine mark\_brooks\_deployment.

1. Also, replace the Engine Label with a label of the form <first\_name>\_<last\_name>, in my case mark\_brooks

A screenshot of a computer

Description automatically generated

→ Click Save

The next dialog looks like this:

A screenshot of a computer

Description automatically generated

→ Click Exit

## Set Permissions on the Deployment

***Important Note: Make sure to complete this step to ensure that no one else will be able to see or run Jobs on your engines!***

* Select the deployment you just created and select the Sharing icon:

A screenshot of a computer

Description automatically generated

* Edit the ACLs so that only you have read, write, and execute permissions (i.e. delete all entries except your own). The edited ACLs in my example look like this:

A screenshot of a computer

Description automatically generated

→ Click Save.

## Start the Deployment

* Select the just-created Deployment and click the Start icon:

A screenshot of a computer

Description automatically generated

* The Deployment state should transition to Active and turn green:

A screenshot of a computer

Description automatically generated

### 

## Generate an Engine Deployment Script

* Select the just-created Deployment and click the Get Install Script item:

A screenshot of a computer

Description automatically generated

* In the dialog that opens, select Run engine in background, Download and Install from Script,and then click the Copy icon:

A screenshot of a computer program

Description automatically generated

## Run the Engine Deployment Script

* In a terminal session, logged in as the admin user, execute this command (in order to pick up the necessary open file limit):

$ sudo su – admin

* Switch to the /streamsets directory:

$ cd /streamsets

Paste in the script copied from the previous step.

When prompted for a download directory, enter the value sdc-download and click enter and then click enter again to confirm.

When prompted for an install directory, enter the value sdc and click enter, and then click enter again to confirm.

Here is an example session with my user input highlighted:

A computer screen with white text

Description automatically generated

If all goes well, after a couple of minutes, you should see an Engine started successfully message like this:

A screen shot of a computer

Description automatically generated

## 

## Confirm the Engine is Healthy

In the StreamSets Platform UI, make sure the engine has a green state and Last Reported Time within the past minute:

A screenshot of a computer

Description automatically generated

# 

# Run a Test Pipeline

To confirm the engine is working, navigate to Build > Pipelines, and create a new blank Data Collector pipeline:

A screenshot of a computer

Description automatically generated

→ Click Next

Confirm that your engine is selected by default:

A screenshot of a computer

Description automatically generated

If it is, click Save & Open in Canvas

If your engine is not selected by default , click the Click here to select link, and select your engine from the list:

A screenshot of a computer

Description automatically generated

Click Save. and then click Save & Open in Canvas.

Add a Dev Data Generator (source) and a Trash (target) to the canvas to construct a trivial pipeline that generates data and then discards it:

A screenshot of a computer

Description automatically generated

Start the pipeline by selecting Start Pipeline from the Draft Run menu:

A screenshot of a computer

Description automatically generated

You should see metrics like this for the running pipeline:

A screenshot of a computer

Description automatically generated

Click the Stop button to stop the pipeline.

Congratulations! Your StreamSets environment is up and running!