

Research Participant Guide

Auto Hadoop Provisioner

Thank you for choosing to participate in this research project. This guide will help you get set up to use the Auto Hadoop Provisioner software. If you need help at any time feel free to reach out to me on teams or through my email: 1803534@uad.ac.uk

There are two methods to getting setup - either using Docker or a Virtual Machine software (VMWare/VirtualBox). Ideally, I would suggest using Docker but if you don't have Docker installed on your computer then go with the Virtual Machine guide.

IMPORTANT Make sure to have completed the consent form before continuing - find link to it [here](#)

Tasks - Time Estimate: ~10min.

After following one of the setup guides below, there are just 5 tasks that I would like you to complete. Each task will have you test a different component of the app. A **survey** will be used to track your experience of this - find the link to it [here](#).

TIP: In order to complete some of these steps it may help to look through the documentation on the homepage of the application

1. Navigate to the *Cluster Builder* tab and start a 3 Node Hadoop Cluster with YARN and Spark enabled.
2. Navigate to the *Cluster Explorer* tab and click on the *Name Node*.
3. Navigate to the *Cluster Monitor* tab and run a job with the following attributes:
 - Job Type - **Spark**
 - Operation - **Pi**
 - Number of Maps - **1000**
 - Map Sample Size - **16**
4. While on the *Cluster Monitor* tab, view job status by clicking the **YARN UI** link. Then navigate back to app.
5. Navigate to *Cluster Builder* and shutdown the cluster.

Docker (Recommended) - Time Estimate: ~5min.

Make sure to have Docker installed - if not see the guide [here](#).

Note Time estimate could be extended if you need to install Docker.

Steps

1. Start Docker
2. Open a terminal/cmd instance
3. Enter these commands to pull & start the docker containers (It may take a minute or two to pull these images)

- `docker run -dp 3000:3000 --name hadoopcreator johnnoon74/hadoopcreator:latest`
- `docker run -dp 5000:5000 -v //var/run/docker.sock:/var/run/docker.sock -v //usr/bin/docker:/usr/bin/docker --name hadoopapi johnnoon74/hadoopapi:latest`

4. Navigate to <http://localhost:3000/>

- Explore the site and when ready follow tasks above

Virtual Machine - Time Estimate: ~15min.

Make sure to have VMWare or VirtualBox installed and updated

DISCLAIMER Since the Virtual Machine will be running Docker in it, overall performance may drop.

Steps

1. Download the VMWare Image [here](#) and Virtual box Image [here](#)
2. Navigate to your VM software and *open* an existing Virtual Machine
3. Startup Linux Virtual Machine
4. Enter login password: **testing**
5. Open the Terminal
6. Ensure the docker containers are running with the command: `docker ps`
 - If containers are not running
 - Enter: `docker restart hadoopcreator`
 - Enter `docker restart hadoopapi`
7. Navigate to <http://localhost:3000/>
 - Explore the site and when ready follow tasks above

Partipant Survey - Time Estimate: ~10min.

After you have finished experimenting with the Hadoop Auto Provisioner App please fill out this **survey**. There are two sections - one for measuring usability (10 questions) and another to get more detailed feedback on what you thought of the app (4 questions).

Find the link to the survery [here](#)