

WEEK 5 REFLECTION

Large-Scale Data Storage Systems – DATA-5400 | Spring 2020

Christina Morgenstern

Installing Hadoop on my VM was relatively straight-forward. But when trying to allocate 8GB of RAM, my computer was complaining. So, I had to specify less, approx. 5 GB.

Also, the preinstalled keyboard layout was American, and I had a hard time trying to find all the relevant keys. You are so accustomed to your keyboard that it gets you into real issues if not only Z and Y are exchanged but backslashes, stars and the like are not where they are supposed to be. At first, I tried to look up the keys on the American layout. Later, I realized that there must be an option to change that and found a command in the documentation that allowed to change keyboard layout.

```
sudo dpkg-reconfigure keyboard-configuration
```

Unfortunately, it didn't work out in my case to change the keyboard configuration as stated (<https://docs.bitnami.com/vmware/faq/configuration/configure-keyboard/>). The problem was that the system couldn't find the required packages. After trying a few things, I stopped and used the American keyboard layout as it was. By then I had noted down all the necessary keys.

Also, the vim editor caused some headache. I am not used to using this editor, so I needed to do some research on how to exit the editor or change values in the file. With the wrong keyboard layout this was quite intense.

Moreover, I wanted to copy commands from the lecture slides to minimize typing errors. However, there seems to be no straight-forward way of copying and pasting commands. I ended up typing everything and making many mistakes at the same time.

Since running Hadoop on my VM didn't quite work out as intended, I tried the "easy option" of creating an HDInsight Cluster on Azure. I have tried Azure before using a 30-days-trial, that's why I needed another Email address to sign up. Luckily, my company address is an Outlook address, so I used it to create a free account.

While creating an HDInsight Cluster seems straight forward, I completely failed in this process. The problem seems to be limits in the free subscription. Whenever, I tried to create the cluster, it wouldn't let me because of limits in core availability. The suggested changes in region, number of nodes and CPU types didn't make any difference. In every case, creation of the cluster was abolished due to limits.

I might have to make some changes to my subscription to be able to create an HDInsight Cluster.

Altogether, this was the first week in which I really struggled to get the assignments to work. Reasons being technical issues, possibly memory restrictions of my laptop, and access restrictions when trying to use the Azure resources.