PathoSense

Installation guide III

Installation Guide Part III - Install and test CloudLink

This procedure explains how to install and test CloudLInk. The CloudLink enables data transfer of sequencing run data to the Pathosense Cloud for subsequent data analysis. This way, manual data transfers are avoided.

Prerequisites

- Installation Guide Part I - Install Ubuntu 22.04 is completed
- Installation Guide Part II - Install MinKNOW and test GPU basecalling is completed

III.I Install docker

1. Open the terminal
2. Execute the following commands to install Docker. Enter “Y” when prompted.

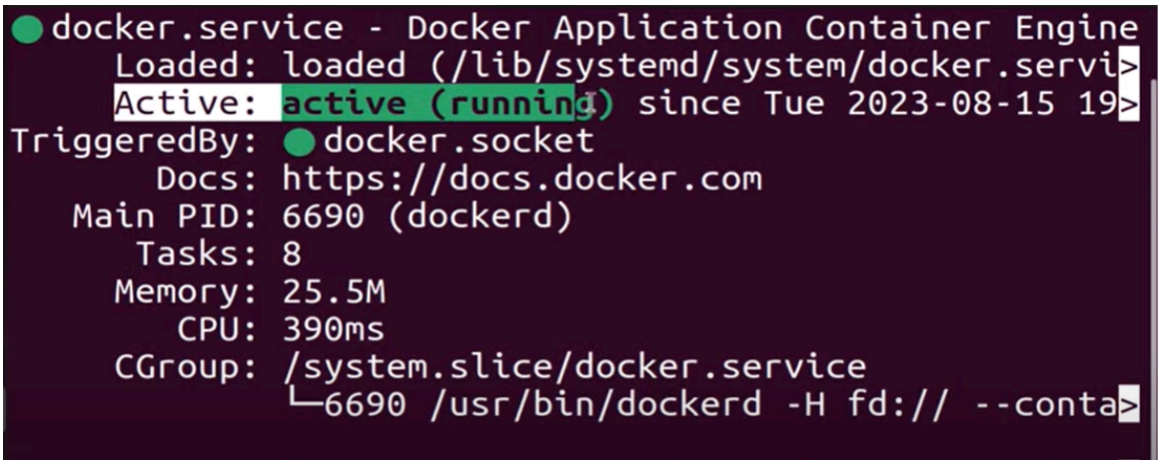
   ```
   sudo apt-get update

   sudo apt install docker.io
   ```

3. Enter the following command to make sure docker starts automatically on start-up

   ```
   sudo systemctl enable docker
   ```

4. Check if docker is active

   ```
   sudo systemctl status docker
   ```



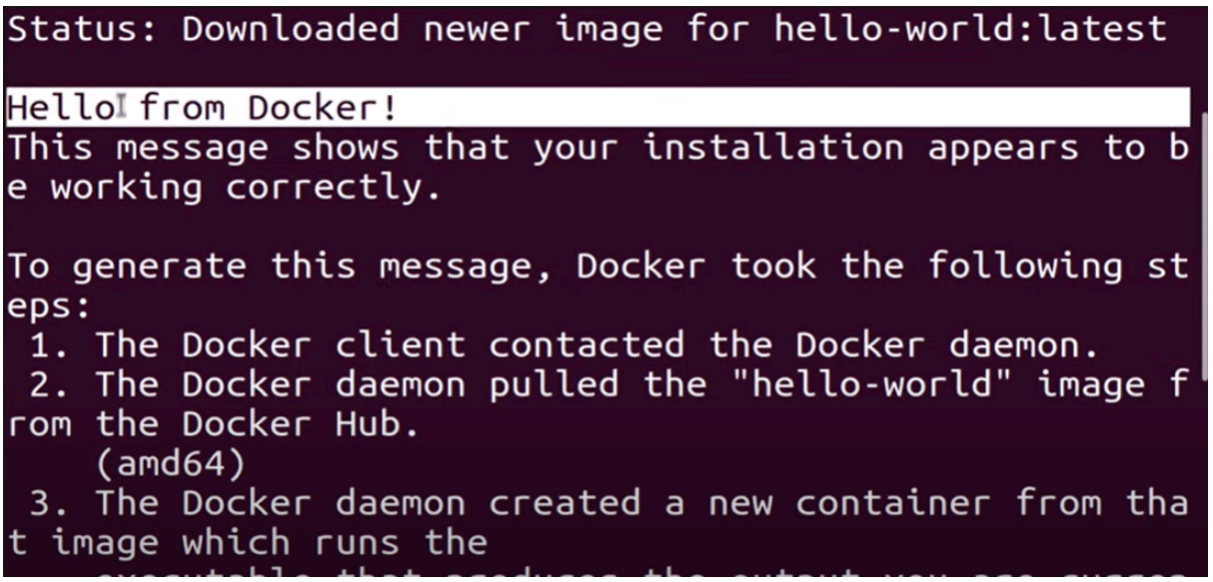
If docker is not yet active, run the following command.

```
sudo systemctl start docker
```

5. Press ctrl + c to return to the command line
6. Test the docker installation by running the following command

```
sudo docker run hello-world
```

If the installation was successful, you will get the ‘Hello from Docker!’’



7. Run the following commands to enable running docker without sudo access

```
sudo groupadd docker
```

```
sudo usermod -aG docker $USER
```

8. Reboot the computer
9. After rebooting, open a terminal. Test if you can run docker without sudo access by running the following command

   ```
   docker run hello-world
   ```

   You should again get the ‘Hello from Docker!’’ message as shown above if everything is set up correctly.

PathoSense

Installation guide III

III.II Install CloudLink

1. Open a terminal
2. Execute the following commands to make a directory named ‘CloudLink’ in your home directory, and move to that directory.

   ```
   cd ~

   mkdir CloudLink

   cd CloudLink
   ```

3. Install the dependencies wget and curl. Press “y” when prompted.

   ```
   sudo apt-get install wget

   sudo apt-get install curl
   ```

4. Run the following command to download the PathoCloud installer.

   ```
   wget https://cloud.pathosense.com/downloads/cloudlink-installer.sh
   ```

5. Run this command to see the content of your current folder.

   ```
   ll
   ```

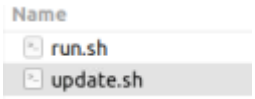
6. Give execution permissions to this installation script. When you run the ‘ll’ command again, you will see the file turned green. This means it is now executable.

   ```
   chmod a+x cloudlink-installer.sh

   ll
   ```

7. Run the installer.
   Enter the login credentials provided by PathoSense.
   Enter the complete absolute path to the ‘PathoSense_Diagnostics’ folder that was chosen as output folder for the sequencing runs in the MinKNOW software.

   ```
   ./cloudlink-installer.sh
   ```

   This will create a *“CloudLink”* directory with the application files as shown below.



8. Check if the docker container is running by running the following command.

   ```
   docker ps
   ```

   If the ‘cloudlink’ is listed in the output, the container is running.

PathoSense

Installation guide III

III.III Test the CloudLink connection

1. In your PathoSense_Diagnostics folder, the CloudLink.log file should have appeared. Check the content of this file This file content should currently look similar to this.

```
2024-02-27 16:08:09,969 INFO: ------------------------
2024-02-27 16:08:09,969 INFO: CloudLink initializing...
2024-02-27 16:08:09,969 INFO: Requesting login...
2024-02-27 16:08:10,300 INFO: Response: OK
2024-02-27 16:08:10,301 INFO: Authenticated as sieglinde.coppens@pathosense.com
2024-02-27 16:08:10,301 INFO: Requesting context...
2024-02-27 16:08:10,744 INFO: Response: OK
2024-02-27 16:08:10,746 INFO: Context loaded: {'lab': 1}
2024-02-27 16:08:10,746 INFO: Watching for new runs...
```

2. Close the file
3. Cut and paste the previously extracted test sequencingrun folder into your home directory (make sure you cut and paste, not copy and paste!)
4. Cut and paste the folder back to the PathoSense_Diagnostics folder. This simulates a new run being added to the folder.
5. Again, check the content of the CloudLInk.log file. The log file should mention the detection of a this ‘new’ run.

```
2024-02-27 16:14:45,432 INFO: Validating...
2024-02-27 16:14:45,440 INFO: -------
2024-02-27 16:14:45,440 INFO: New run: /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingrun/20230222_
2024-02-27 16:14:45,441 INFO: Validation successful.
2024-02-27 16:14:45,441 INFO: Analyzing directory /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingru
2024-02-27 16:14:45,453 INFO: Parsing run file...
2024-02-27 16:14:45,643 INFO: Uploading outputs from /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencin
2024-02-27 16:14:45,644 INFO: Uploading /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingrun/2023022
report_FAV65050_20230222_0922_85c46a1a.html to 1/2024-02-27T16:14:45.441890_100/report_FAV65050_20230222_0922_85c46a1a.html
2024-02-27 16:14:46,083 INFO: Uploading /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingrun/2023022
240227_barcode67_1-0004176_unknown.fastq.gz to 1/2024-02-27T16:14:45.441890_100/240227_barcode67_1-0004176_unknown.fastq.gz
2024-02-27 16:14:46,245 INFO: Uploading /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingrun/2023022
240227_barcode68_1-0004177_unknown.fastq.gz to 1/2024-02-27T16:14:45.441890_100/240227_barcode68_1-0004177_unknown.fastq.gz
2024-02-27 16:14:46,377 INFO: Uploading /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingrun/2023022
240227_barcode67_1-0004176_unknown_fast5.tar.gz to 1/2024-02-27T16:14:45.441890_100/240227_barcode67_1-0004176_unknown_fast5.ta
2024-02-27 16:14:46,640 INFO: Uploading /Test_sequencingrun-20240227T161228Z-001/Test_sequencingrun/Test_sequencingrun/2023022
240227_barcode68_1-0004177_unknown_fast5.tar.gz to 1/2024-02-27T16:14:45.441890_100/240227_barcode68_1-0004177_unknown_fast5.ta
2024-02-27 16:14:46,833 INFO: Calling home...
2024-02-27 16:14:46,833 INFO: Requesting nextflow...
2024-02-27 16:14:47,127 INFO: Response: OK
2024-02-27 16:14:47,128 INFO: False
```

6. Email info@pathosense.com to verify the data was successfully received in the cloud.

The CloudLink installation is now complete. Data from sequencing runs that are outputted in the PathoSense_Diagnostics folder will now automatically be uploaded to and analyzed in the PathoSense cloud environment.