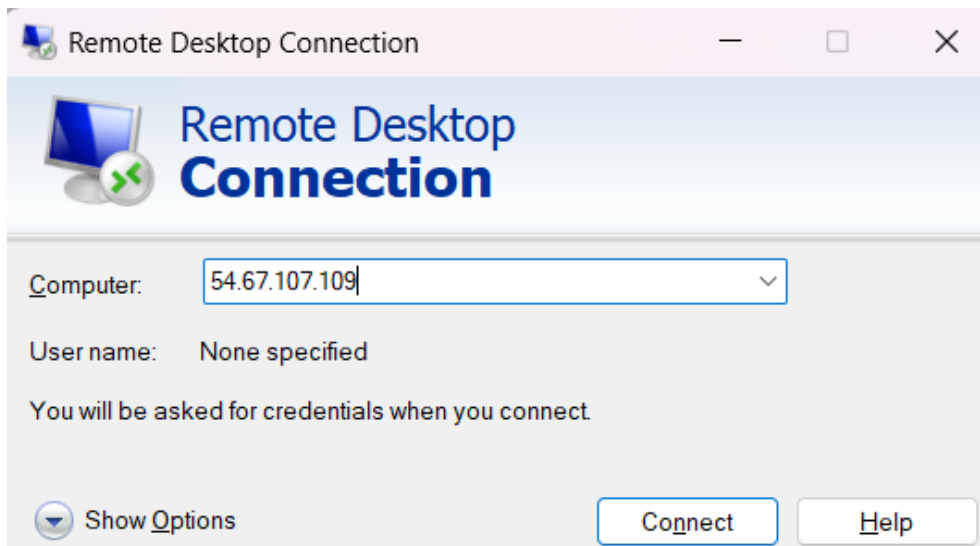


Apache Flink Getting Started

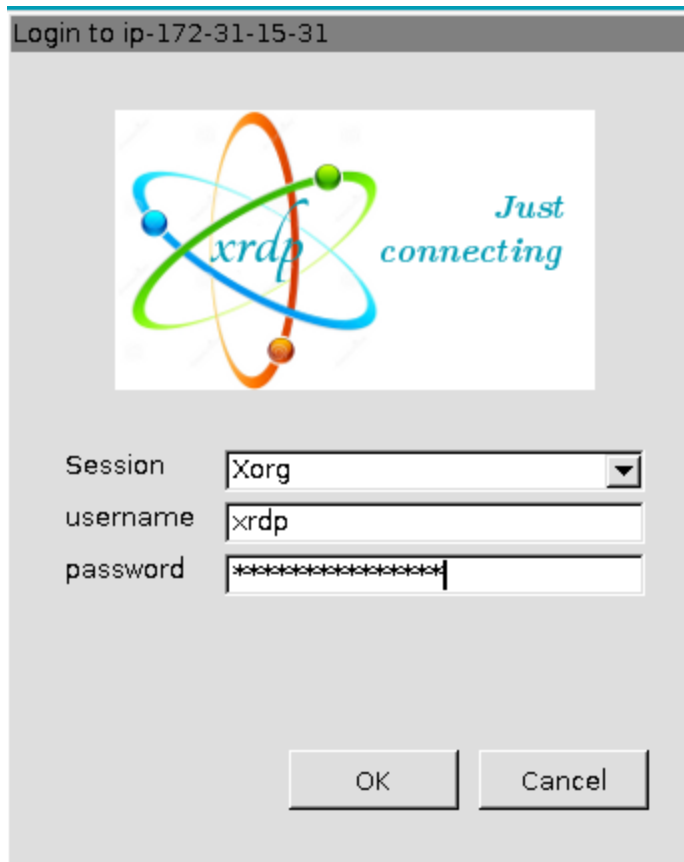
Experiment 1: Hello World

1.1 Steps to run your first Flink Program

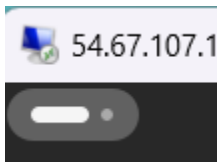
- 1.1.1 Browse to the GitHub repo that you cloned. This should be cloned to your Windows Jump Box and the Flink Development Server
<https://github.com/GeorgeNiece/flink-data-processing-2day>
- 1.1.2 From a command prompt on your jumpbox machine SSH to the Ubuntu server
ssh -o ServerAliveInterval=180 -o ServerAliveCountMax=2 -i ansible.pem ubuntu@ip_address_provided
- 1.1.4 Change to the flink folder, verify Flink isn't started, start the Flink dev cluster, and verify that it started
ps -ef | grep flink
cd ~/flink-2.0.0
./bin/start-cluster.sh
ps -ef | grep flink



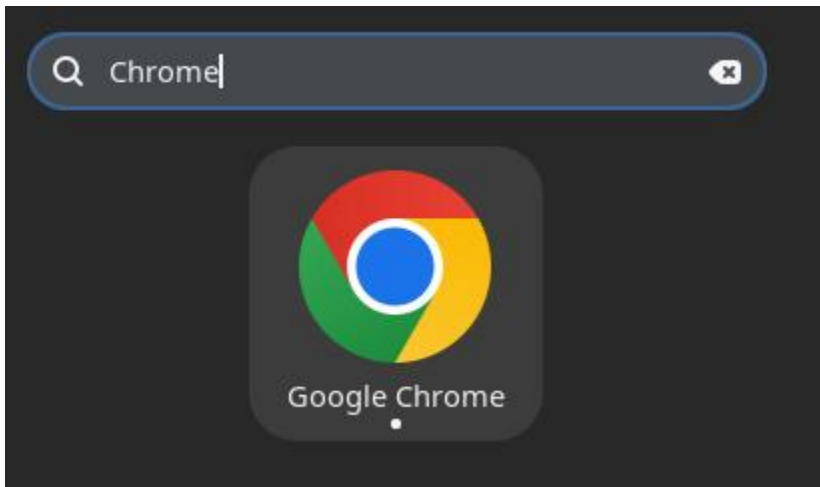
- 1.1.5 Login to the ubuntu dev sandbox using Windows RDP with the xrdp user and the password that you set in Step 1.14



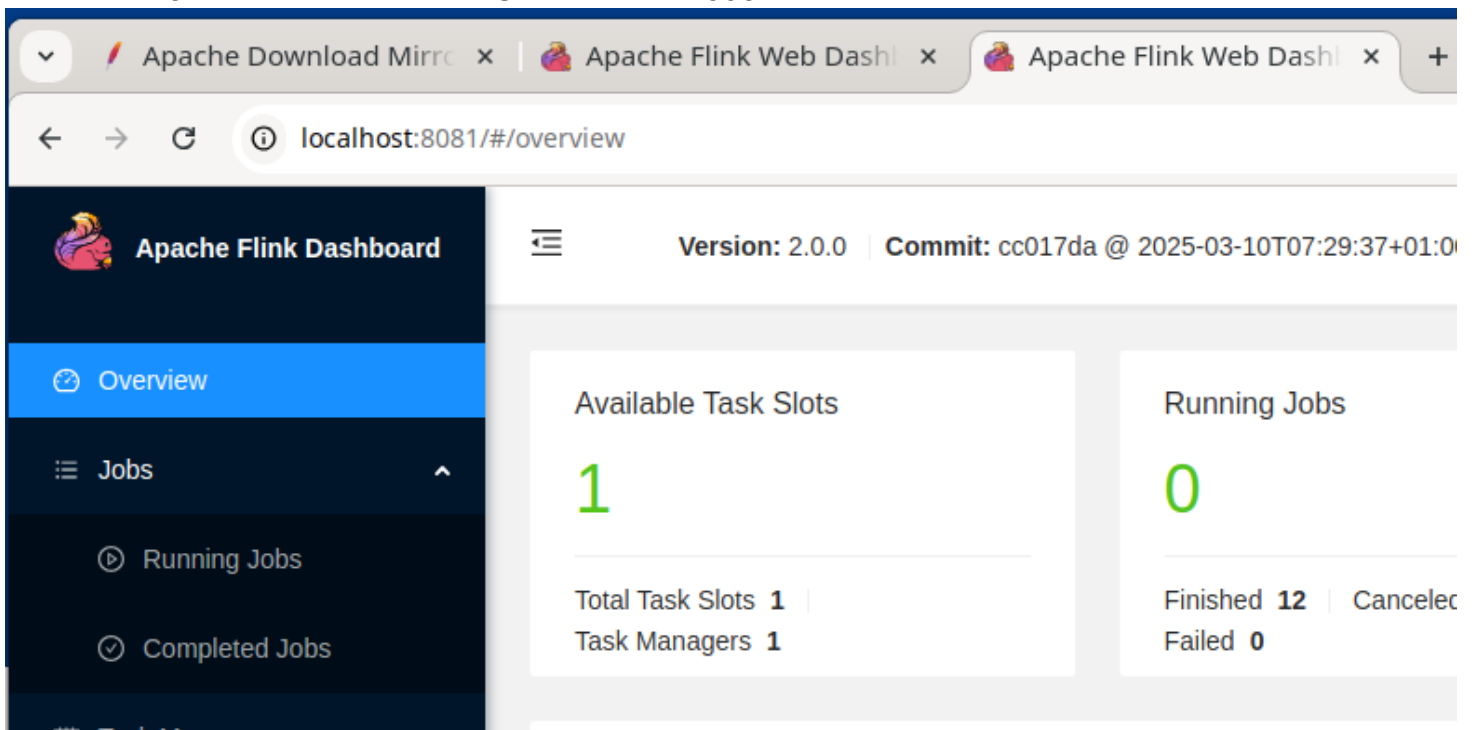
1.1.6 Click the Activities button in the top left corner of the Ubuntu Desktop



1.1.7 Wait for the Search Box at the top of the Ubuntu Desktop, and enter Chrome, click on the Launch Logo



1.1.8 Load the Flink Web UI at localhost:8081



1.1.9 Click the Job Manager in the left hand navigation

Apache Flink Dashboard

Overview

Jobs

Running Jobs

Completed Jobs

Task Managers

Job Manager

Version: 2.0.0 | Commit: cc017da @ 2025-03-10T07:29:37+01:00

Metrics Configuration Logs Stdout Log List Thread Dump

Flink Memory Model

Effective Configuration

Total Process Memory	JVM Heap	1.00 GB
Total Flink Memory	Off-Heap Memory	128 MB
JVM Heap		
Off-Heap		

1.1.10 Select Log List in the page navigation. These are the logs we'll monitor while we're running some of our experiments

Apache Flink Dashboard

Version: 2.0.0 | Commit: cc017da @ 2025-03-10T07:29:37+01:00 | Mess

Metrics Configuration Logs Stdout **Log List** Thread Dump Profi

Log Name	Last Modified Time	Size (KB)
flink-ubuntu-taskexecutor-0-ip-172-31-15-31.out	2025-04-13 22:08:34.730	0.36
flink-ubuntu-taskexecutor-0-ip-172-31-15-31.log	2025-04-14 00:05:49.716	164.46
flink-ubuntu-taskexecutor-1-ip-172-31-15-31.log.1	2025-04-13 16:07:54.038	49.7

1.1.11 The first `flink-ubuntu-taskexecutor*.out` file will be the one we spend the most time looking at. Select that so that we can view there. We could open two browser tabs and watch the Jobs -> Running Jobs page while we run our first experiment.

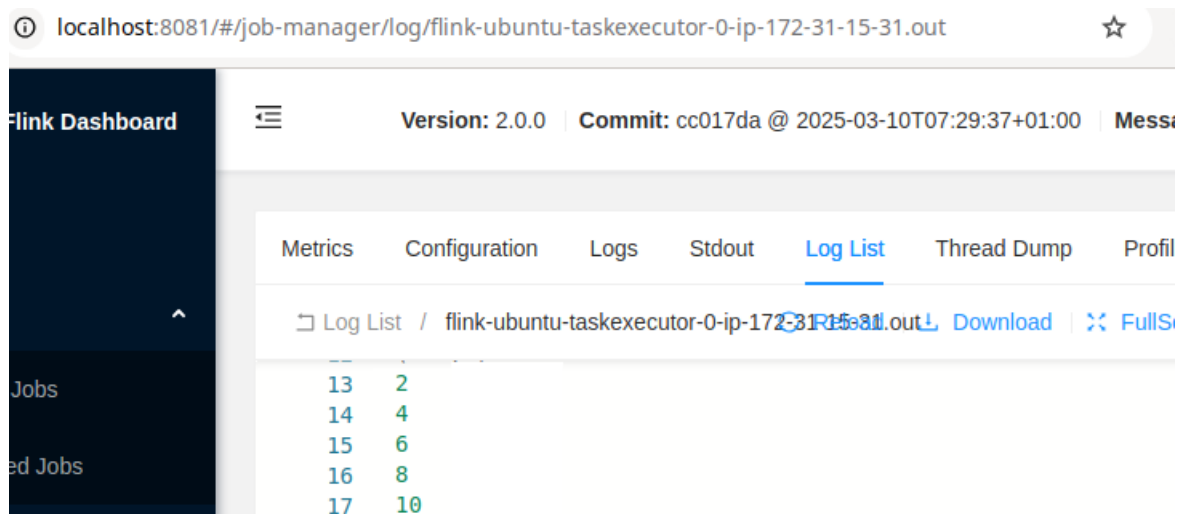
1.1.12 Navigate back to the SSH terminal to the flink distribution folder

```
cd ~/flink-2.0.0
```

```
./bin/flink run ~/flink-data-processing-2day/experiments/built/HelloWorld.jar
```

This should show us with the Job submission, Program execution finished, JobID and the Job Runtime

View the first **`flink-ubuntu-taskexecutor*.out`**



1.1.13 **Congratulations, time to celebrate** you ran your first Flink program in our session