

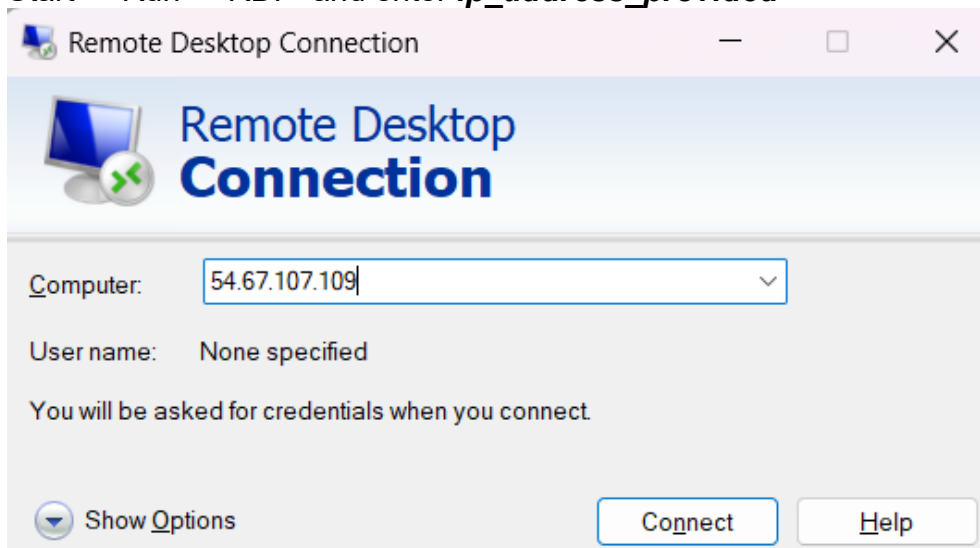
# Apache Flink Getting Started

## Experiment 0: Setup

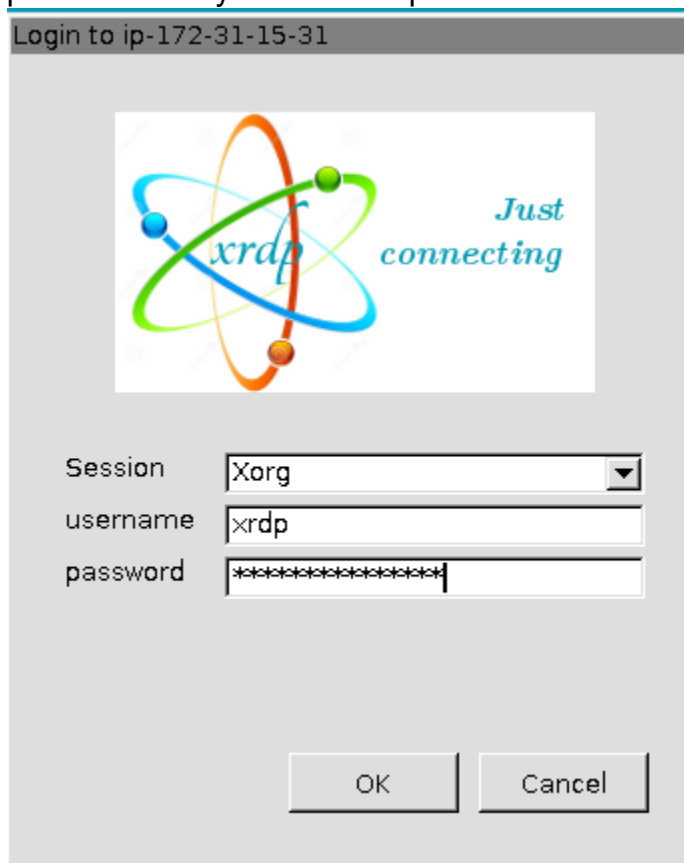
### 1.1 Steps to setup your Experiment Account

- 1.1.1 Use the provided information to connect to your Windows Jump box environment. These will be in the GitHub and the password will be shared in chat for those.
- 1.1.2 Grab your flink sandbox credentials to connect, IP address, user (ubuntu), and pem file. This is in the GitHub Repo provided for this class (<https://github.com/GeorgeNiece/flink-data-processing-2day>), and the FlinkDev-2025-04-14 excel file shared by the Flink course leader. You can copy the pem file after downloading and paste to the desktop in the Windows Jump Box
- 1.1.3 Ensure the pem file is in the same folder that you do the ssh from following and that you've changed the permissions for the file  
In Windows  
**attrib -x ansible.pem**  
In Mac  
**chmod 600 ansible.pem**
- 1.1.4 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 the Server RDP user password, (make sure to keep this for access to the Flink WebUI through Remote Desktop Protocol)  
**sudo passwd xrdp**
- 1.1.5 Retrieve the Flink download we need with  
**wget https://dlcdn.apache.org/flink/flink-2.0.0/flink-2.0.0-bin-scala\_2.12.tgz**
- 1.1.6 Unpack that in your ubuntu user home folder  
**cd ~**  
**tar -xzf flink-2.0.0-bin-scala\_2.12.tgz**
- 1.1.7 Download the GitHub repo on the dev server  
**cd ~**  
**git clone https://github.com/GeorgeNiece/flink-data-processing-2day**
- 1.1.8 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**

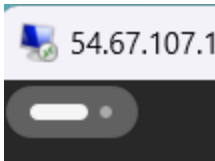
- 1.1.9 From the Windows Jump Host, RDP to the dev server  
Start -> Run -> RDP and enter ***ip\_address\_provided***



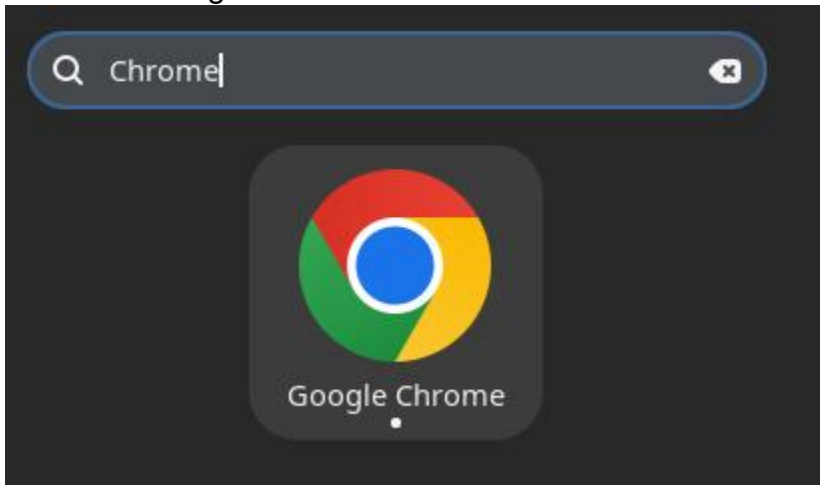
- 1.1.10 Login to the ubuntu dev sandbox using Windows RDP with the xrdp user and the password that you set in Step 1.14



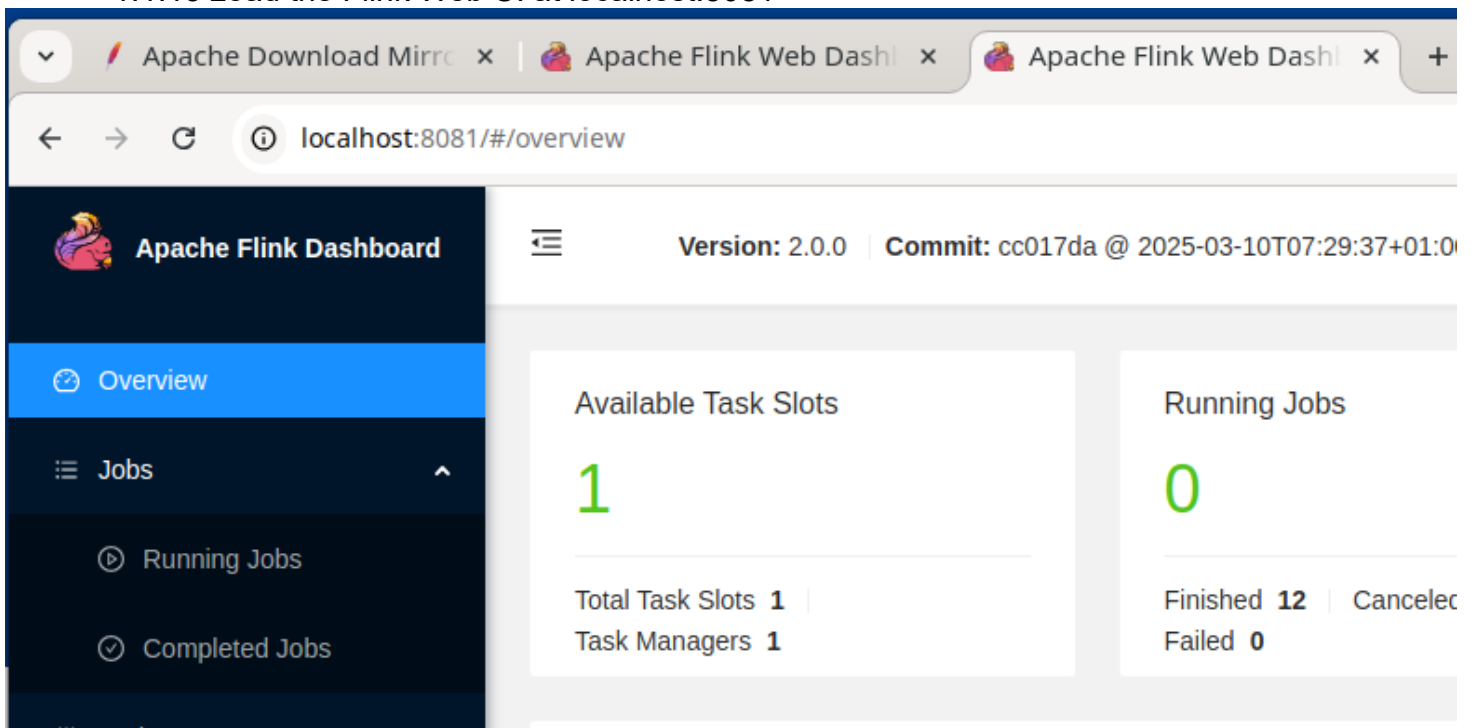
- 1.1.11 Click the Activities button in the top left corner of the Ubuntu Desktop



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



1.1.13 Load the Flink Web UI at localhost:8081



1.1.14 Click the Job Manager in the left hand navigation

The screenshot shows the Apache Flink Dashboard at `localhost:8081/#/job-manager/metrics`. The sidebar on the left contains the following navigation items:

- Overview
- Jobs
  - Running Jobs
  - Completed Jobs
- Task Managers
- Job Manager (highlighted)

The main content area is titled "Metrics" and includes tabs for Configuration, Logs, Stdout, Log List, and Thread Dump. The "Metrics" tab is active, displaying the "Flink Memory Model" diagram and a table of "Effective Configuration".

**Flink Memory Model Diagram:**

```

graph TD
    subgraph Total_Process_Memory [Total Process Memory]
        subgraph Total_Flink_Memory [Total Flink Memory]
            JVM_Heap[JVM Heap]
            Off_Heap[Off-Heap]
        end
    end
  
```

**Effective Configuration Table:**

Configuration Item	Value
JVM Heap	1.00 GB
Off-Heap Memory	128 MB

1.1.15 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)
<a href="#">flink-ubuntu-taskexecutor-0-ip-172-31-15-31.out</a>	2025-04-13 22:08:34.730	0.36
<a href="#">flink-ubuntu-taskexecutor-0-ip-172-31-15-31.log</a>	2025-04-14 00:05:49.716	164.46
<a href="#">flink-ubuntu-taskexecutor-1-ip-172-31-15-31.log.1</a>	2025-04-13 16:07:54.038	49.7

1.1.16 The first `flink-ubuntu-taskexecutor*.out` file will be the one we spend the most time looking at.

1.1.17 Navigate back to the SSH terminal and open the Flink configuration

```
cd ~/flink-2.0.0
./vi conf/config.yaml
```

1.1.18 Review the file (if you're not VI proficient, use "`nano conf/config.yaml`")

1.1.19 **Congratulations, time to celebrate** finishing the setup of a Flink Developer Sandbox