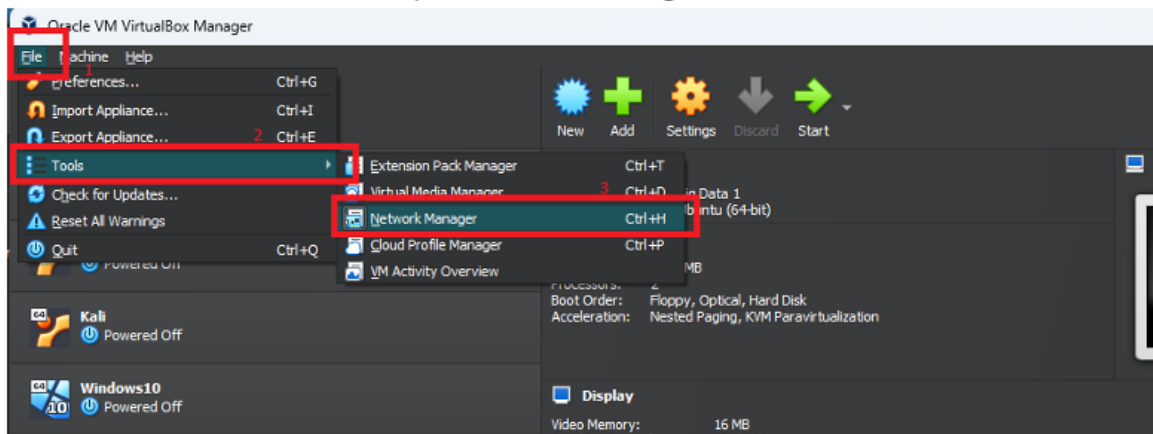
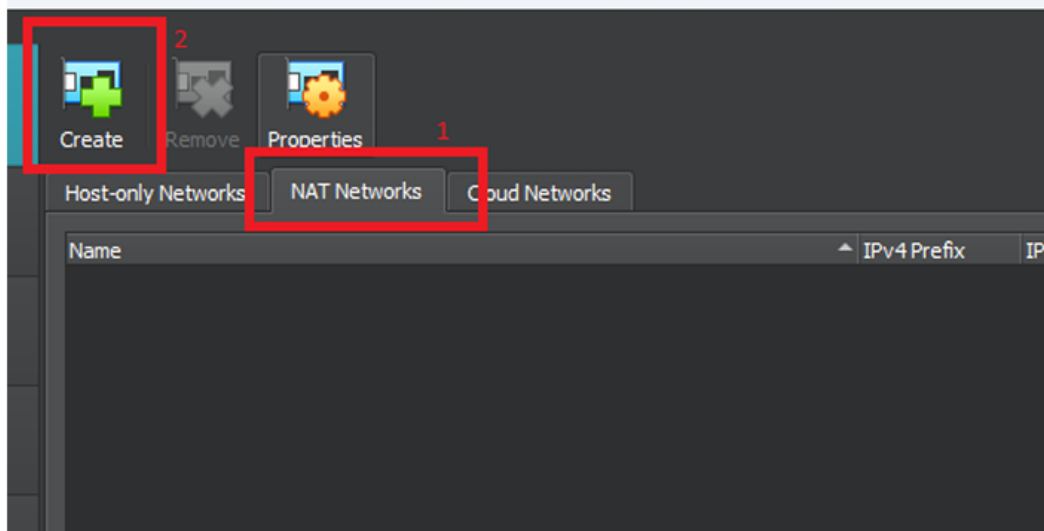


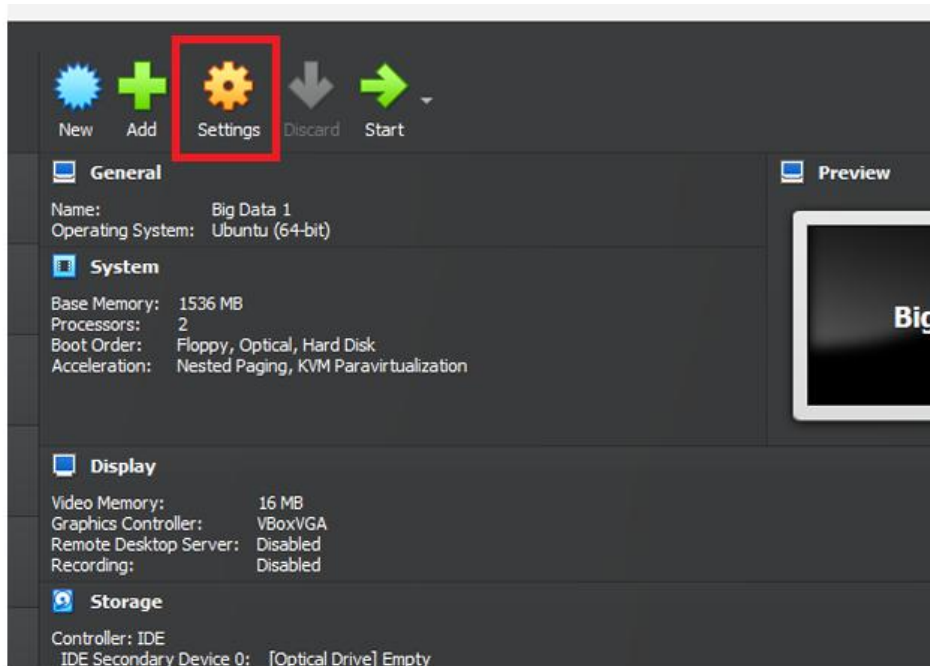
- i. Please launch VirtualBox and open Network Manager:



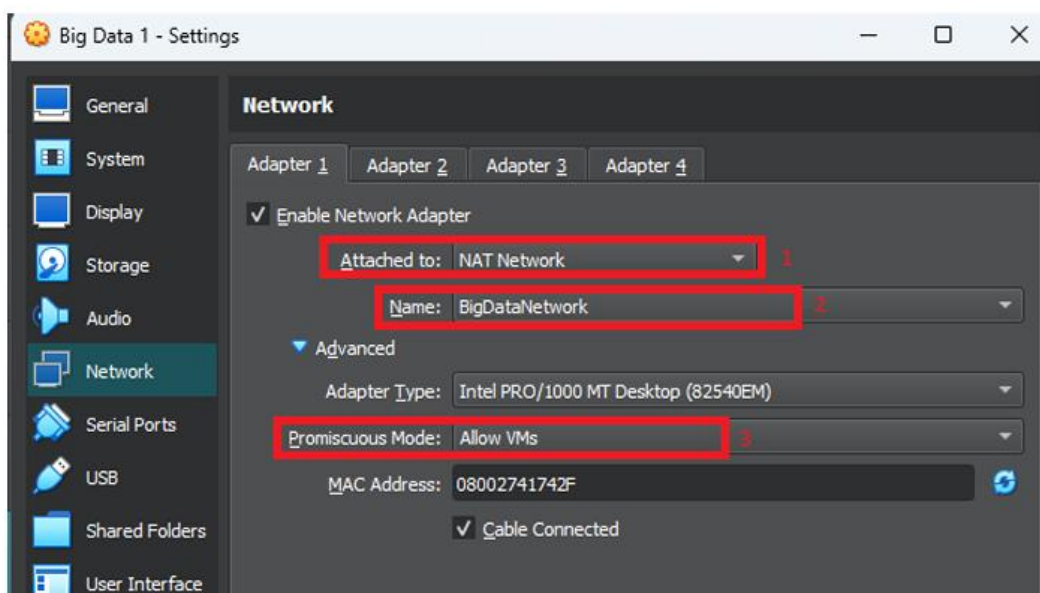
- ii. Then select NAT Networks, and press create... This will give you a NAT Network called 'NatNetwork'



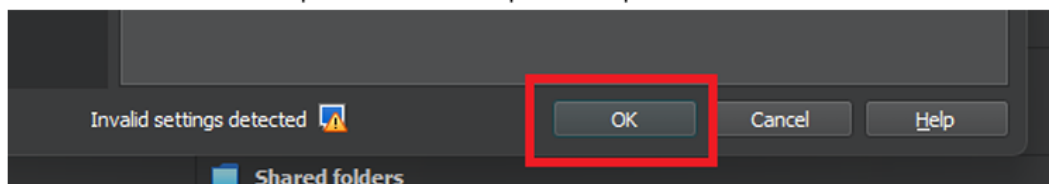
- iii. Then for each of your VMs you need to click on Settings,



- iv. Then go to network, then adapter 1, then Attached to... Set this to 'NAT Network', and check that allow VMs is checked under promiscuous mode:

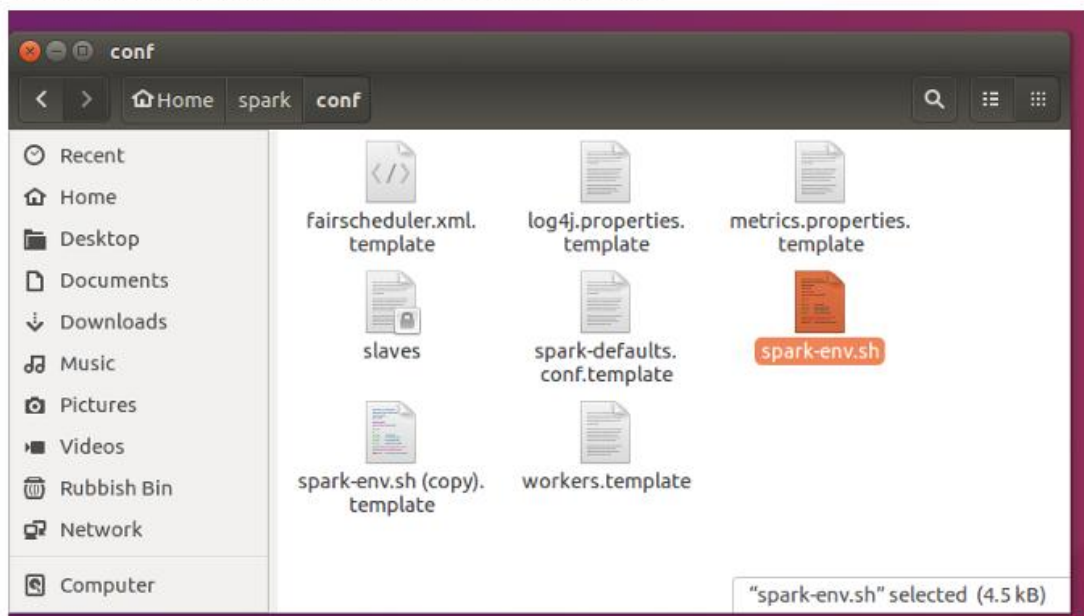


- v. Then click on OK and repeat the last couple of steps for the other VM.



Part 2: Spark cluster launching

- i. For two VMs, open terminals and use
ifconfig
to get the IP address for each VM
- ii. Use
sudo gedit /etc/hosts
command to open the hosts file and change the IP addresses for both m1 and m2, for BOTH virtual machines
- iii. In m1, find the file—spark-env.sh (as below) and change the IP address at the next to the bottom line to the IP address of m1



- iv. Direct to the spark folder using
cd spark
and initializing the spark cluster consisting of m1 and m2 by executing:
./sbin/start-all.sh
- v. Now please open the web browser at m1 and enter the following:
<http://m1:8080/>
to view the Spark cluster

```
>>>pyspark --master spark://:7077
```