Usage instructions:

1. Launch the product via 1-click from AWS Marketplace. Wait until the instance status changes to 'Running' and passes all health checks. Then, connect to your instance using your Amazon private key and the 'ubuntu' user."

To update software, use: **sudo apt update && sudo apt upgrade -y**

Connect & basics

- SSH: ssh -i <your-key>.pem ubuntu@<PUBLIC_IP>
- Hadoop is installed at /opt/hadoop
- Configs: /opt/hadoop/etc/hadoop
- Logs: /opt/hadoop/logs

On first boot:

- A one-time formatter prepares HDFS.
- Systemd services start daemons directly (no SSH fan-out).

Verify services

 # Check systemd services

systemctl status --no-pager hadoop-dfs hadoop-yarn hadoop-historyserver

 # Check processes and ports

ps -ef | egrep 'NameNode|DataNode|ResourceManager|NodeManager' | egrep -v egrep

ss -ltnp | egrep '(:9870|:9864|:8088|:8042|:19888|:9000)'

Web Uis: in a browser goto:

- NameNode: http://Your_Public_IP_:9870
- DataNode: http://Your_Public_IP_:9864
- YARN RM: http://Your_Public_IP_:8088
- NodeManager: http://Your_Public_IP_:8042
- MR JobHistory: http://Your_Public_IP_:19888

 # Stop stack (order matters)

sudo systemctl stop hadoop-historyserver hadoop-yarn hadoop-dfs

 # Start stack

sudo systemctl start hadoop-dfs hadoop-yarn hadoop-historyserver

 # Restart (example)

sudo systemctl restart hadoop-dfs

AWS Data

- Data Encryption Configuration: This solution does not encrypt data within the running instance.
- User Credentials are stored: /root/.ssh/authorized_keys & /home/ubuntu/.ssh/authorized_keys
- Monitor the health:
    - Navigate to your Amazon EC2 console and verify that you're in the correct region.
    - Choose Instance and select your launched instance.
    - Select the server to display your metadata page and choose the Status checks tab at the bottom of the page to review if your status checks passed or failed.

Extra Information: **(Optional)**

Allocate Elastic IP

To ensure that your instance **keeps its IP during restarts** that might happen, configure an Elastic IP. From the EC2 console:

1. Select ELASTIC IPs.
2. Click on the ALLOCATE ELASTIC IP ADDRESS.
3. Select the default (Amazon pool of IPv4 addresses) and click on ALLOCATE.
4. From the ACTIONS pull down, select ASSOCIATE ELASTIC IP ADDRESS.
5. In the box that comes up, note down the Elastic IP Address, which will be needed when you configure your DNS.
6. In the search box under INSTANCE, click and find your INSTANCE ID and then click ASSOCIATE.
7. Your instance now has an elastic IP associated with it.
8. For additional help: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/elastic-ip-addresses-eip.html