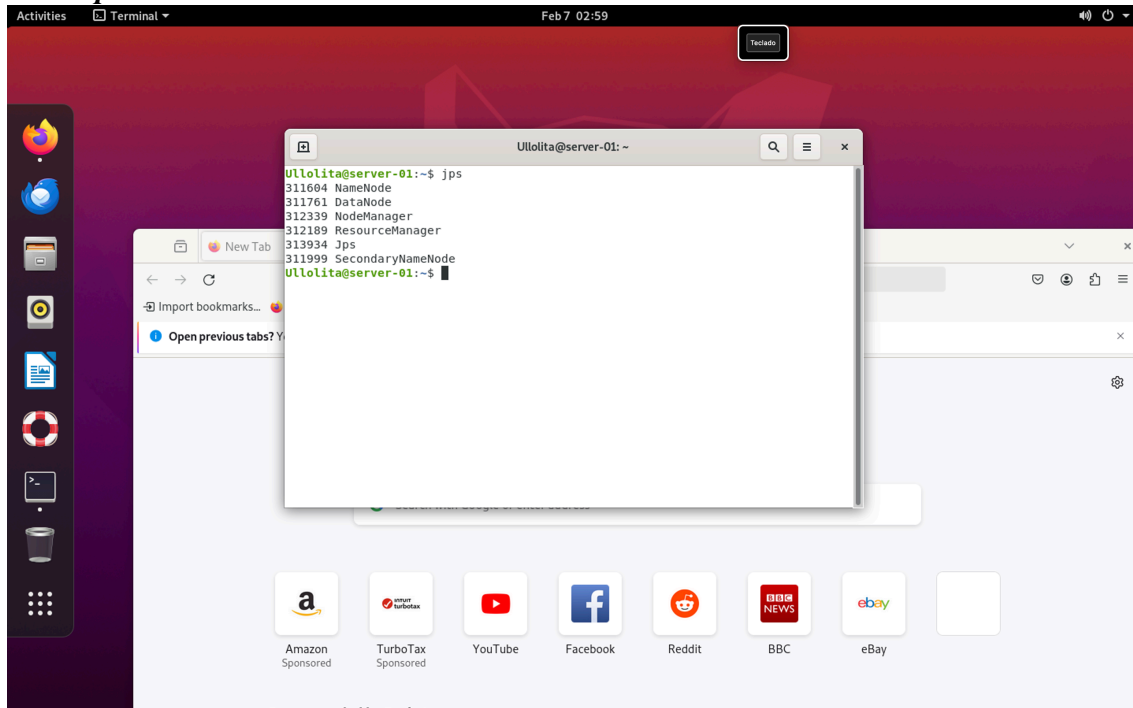


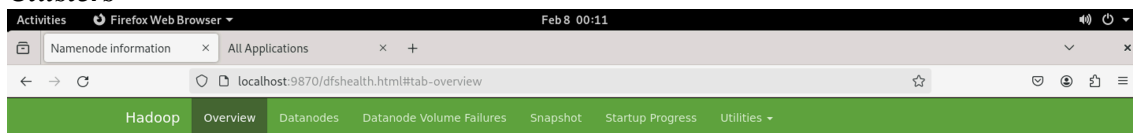
Nombre: Jade Ullola Castro		Matrícula: 4500985
Nombre del curso: Infraestructura para BigData	Nombre del profesor: Miguel de Jesús Martínez Felipe	
Módulo: 1	Actividad: Actividad 4	
Fecha: Febrero 7, 2024		

Actividad 4

Hadoop instalado



Clusters



Overview 'localhost:9000' (✓active)

Started:	Wed Feb 07 23:41:36 +0000 2024
Version:	3.3.6, r1be78238728da9266a4f88195058f08fd012bf9c
Compiled:	Sun Jun 18 08:22:00 +0000 2023 by ubuntu from (HEAD detached at release-3.3.6-RC1)
Cluster ID:	CID-36a6d2a1-966d-43d9-bfee-2a880bd11412
Block Pool ID:	BP-251883396-10.0.0.4-1707348894517

Summary

Security is off.

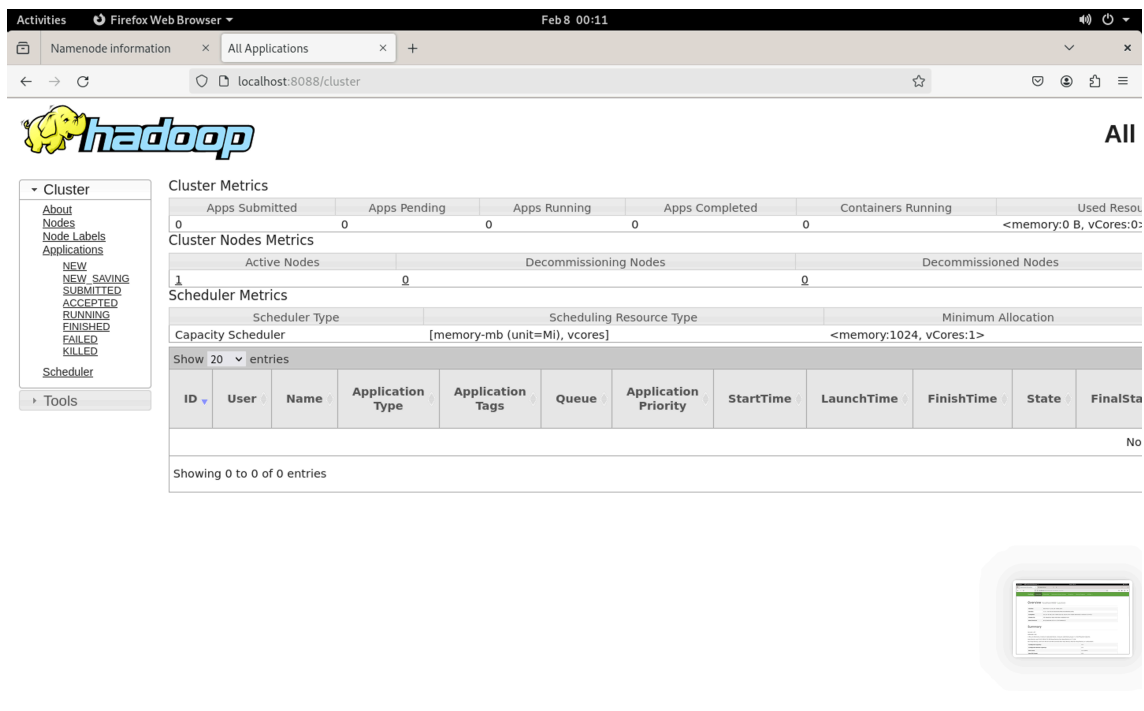
Safemode is off.

1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).

Heap Memory used 124.53 MB of 335 MB Heap Memory. Max Heap Memory is 1.71 GB.

Non Heap Memory used 49.42 MB of 50.84 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	0 B
Configured Remote Capacity:	0 B
DFS Used:	0 B (100%)
Non DFS Used:	0 B



Reflexion

Una vez que la máquina virtual está en funcionamiento, se establece una conexión SSH desde la terminal de la máquina local, en este caso, desde un entorno Mac. Esto se logra utilizando el comando ssh seguido de las credenciales y la dirección IP de la máquina virtual. La instalación de Hadoop en la máquina virtual, llevada a cabo desde la terminal fue un poco complicado debido a la inestable conexión de internet. Sin embargo, no fue algo que no se pudiera solucionar.

Una vez teniendo una conexión estable se tuvo que conectar desde Visual Studio Code con la máquina virtual y trabajar desde la terminal. Acabando con la instalación y la verificación de esta con 'jps'. Una vez configurado, se inician los servicios de Hadoop utilizando los scripts proporcionados. Esto puede incluir servicios como el NameNode, DataNode, ResourceManager y NodeManager. Se verifica que los servicios estén en funcionamiento correctamente.

Con los servicios en marcha, se accede a las interfaces web de Hadoop a través de HTTP. Estas interfaces proporcionan información detallada sobre el estado del clúster, el uso de recursos y el progreso de los trabajos. Se accede a estas interfaces desde el navegador web de la máquina local, ingresando la dirección IP y el puerto asociado a cada interfaz. Una vez dentro de las interfaces, se explora la información disponible y se monitorea el rendimiento del clúster. Esto puede incluir la visualización de estadísticas de nodos, la ejecución de trabajos MapReduce y la supervisión de la replicación de datos.