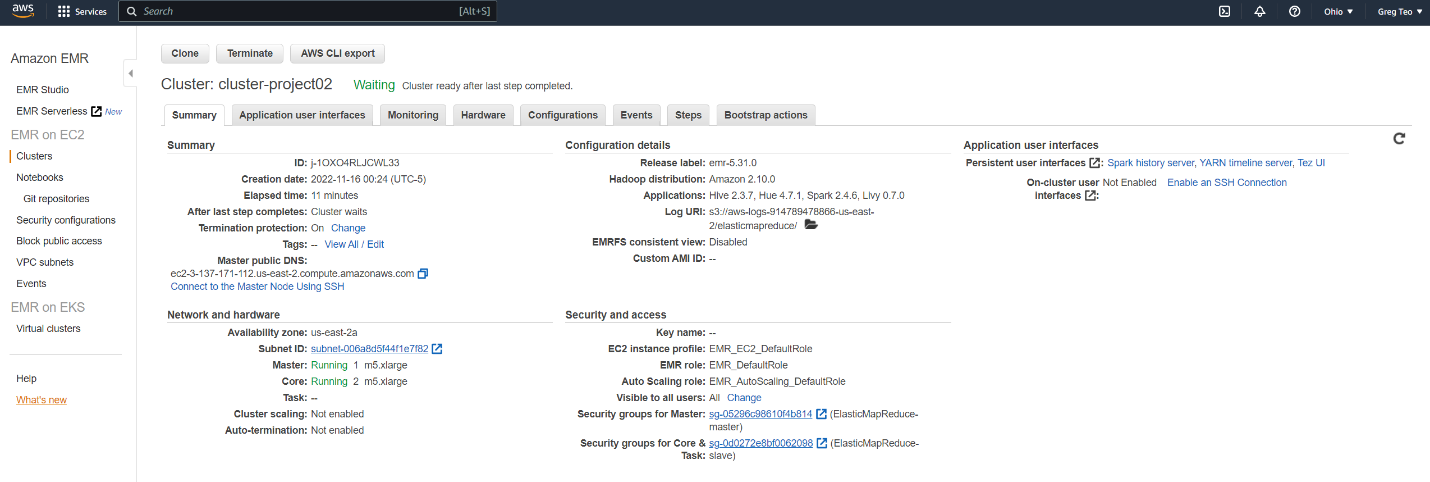
**Overview**

This project uses a Spark cluster on AWS EMR (Elastic Map Reduce), a leading cloud big data platform, to examine fundamental questions about an IMDB dataset. PySpark will be used in this project to carry out basic analysis tasks.

**Cluster Configuration**



**Notebook Configuration**

