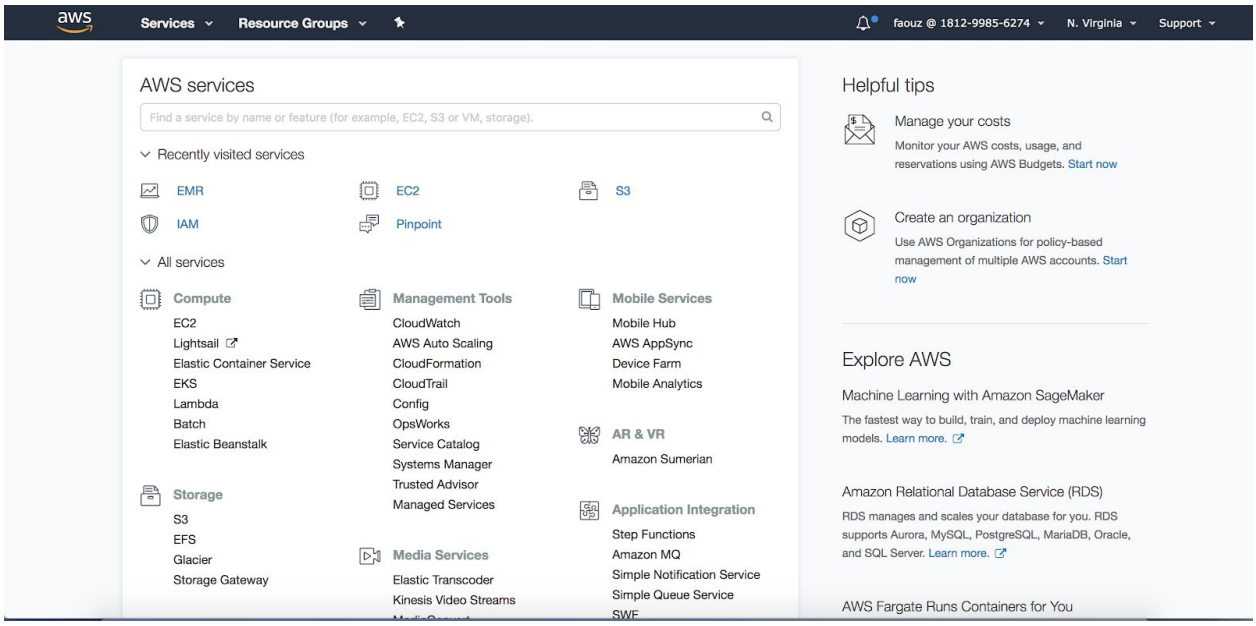
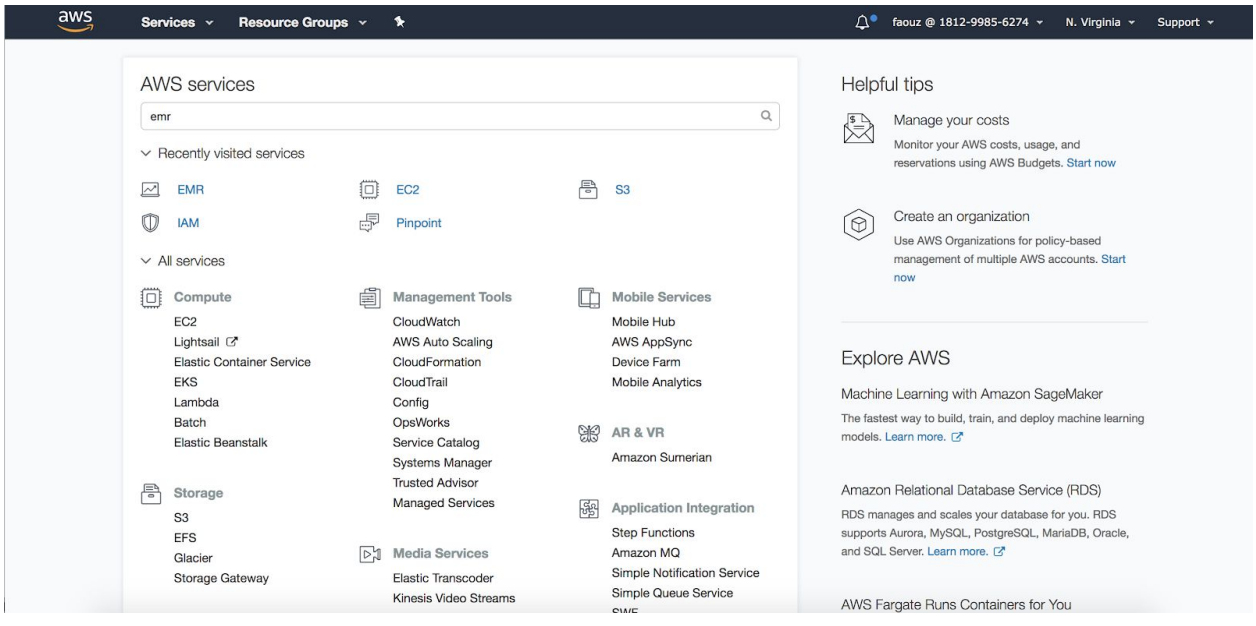
Tutoriel : Créer une cluster Spark 2.3 en 5 minutes sur Amazon Web Services

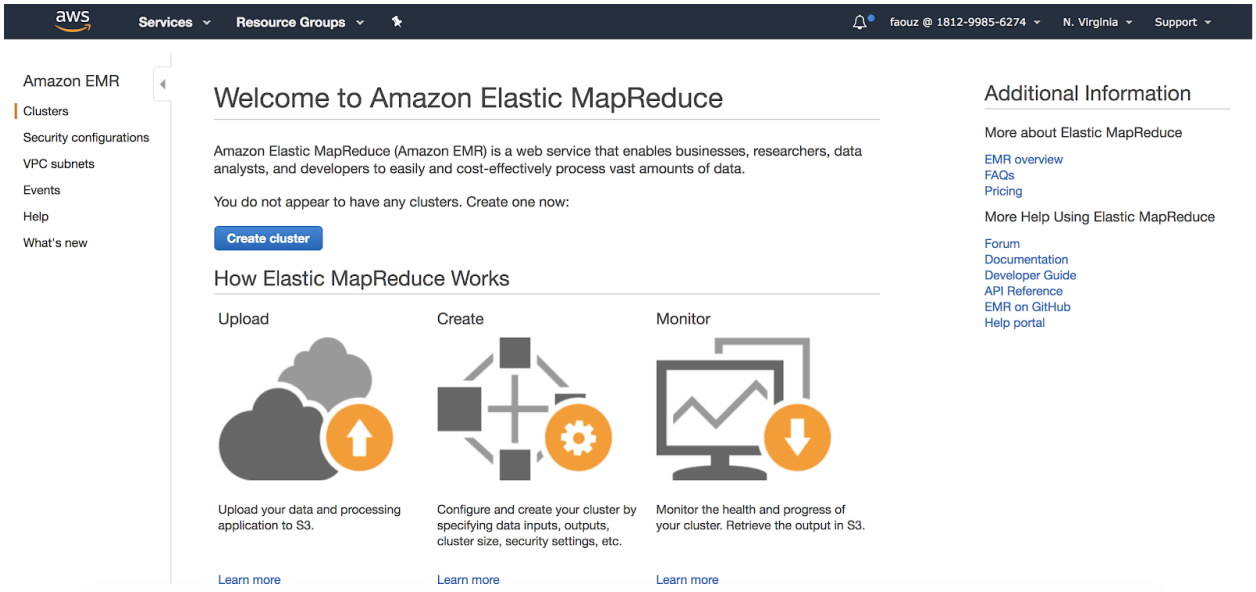
- Connectez-vous sur la console AWS via votre compte Rosetta Hub.



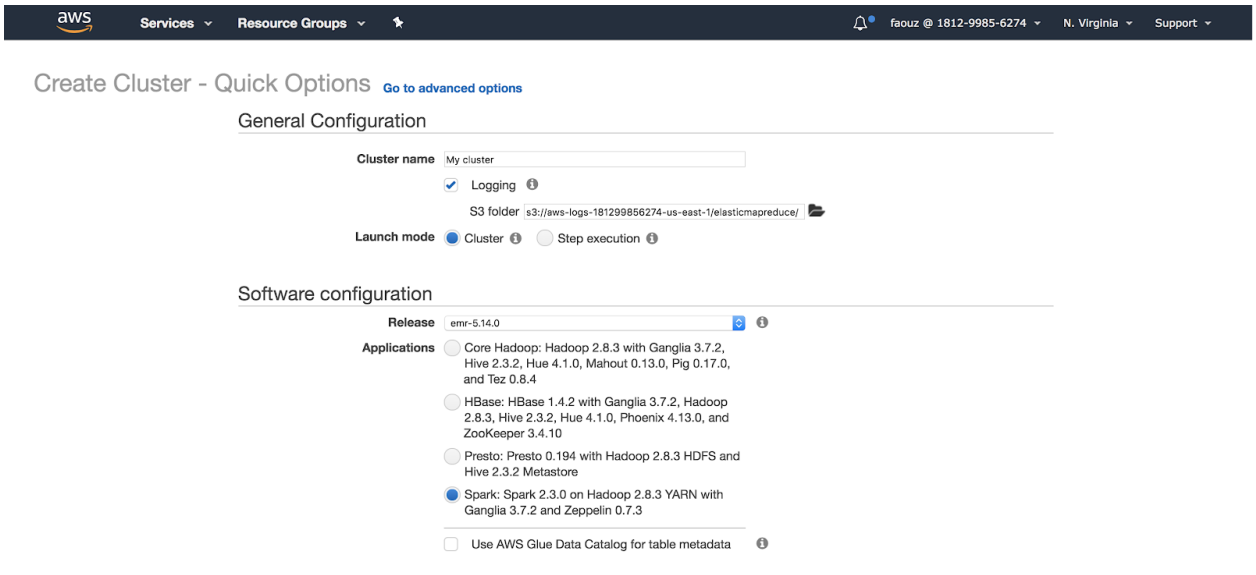
- Recherchez EMR (Elastic MapReduce)



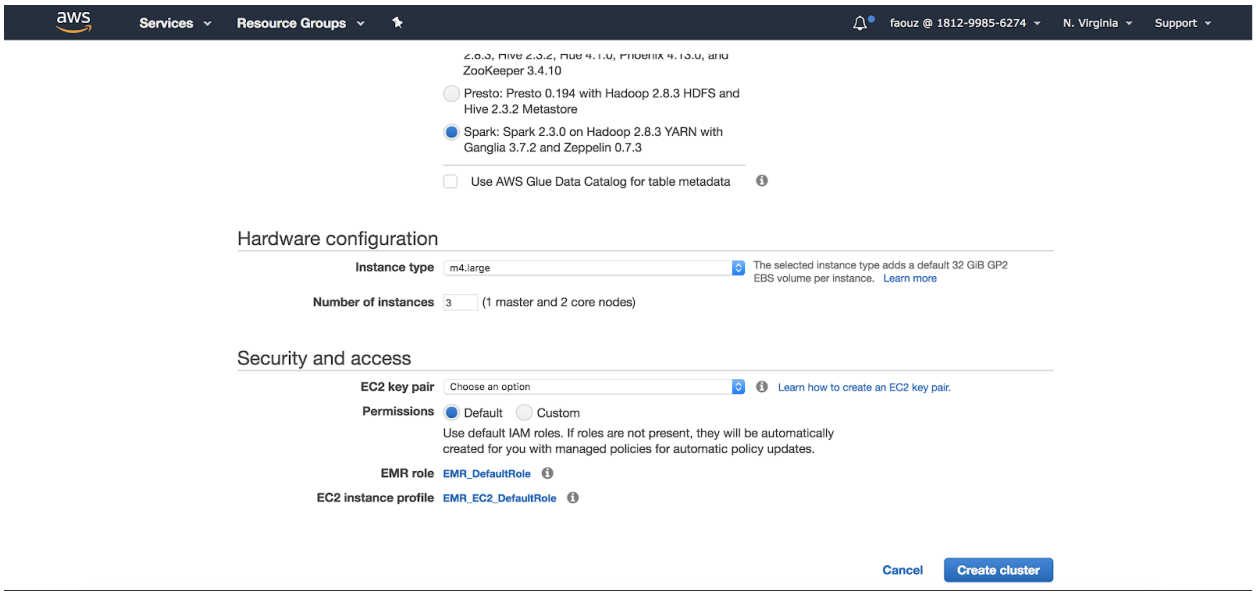
- Créez un cluster



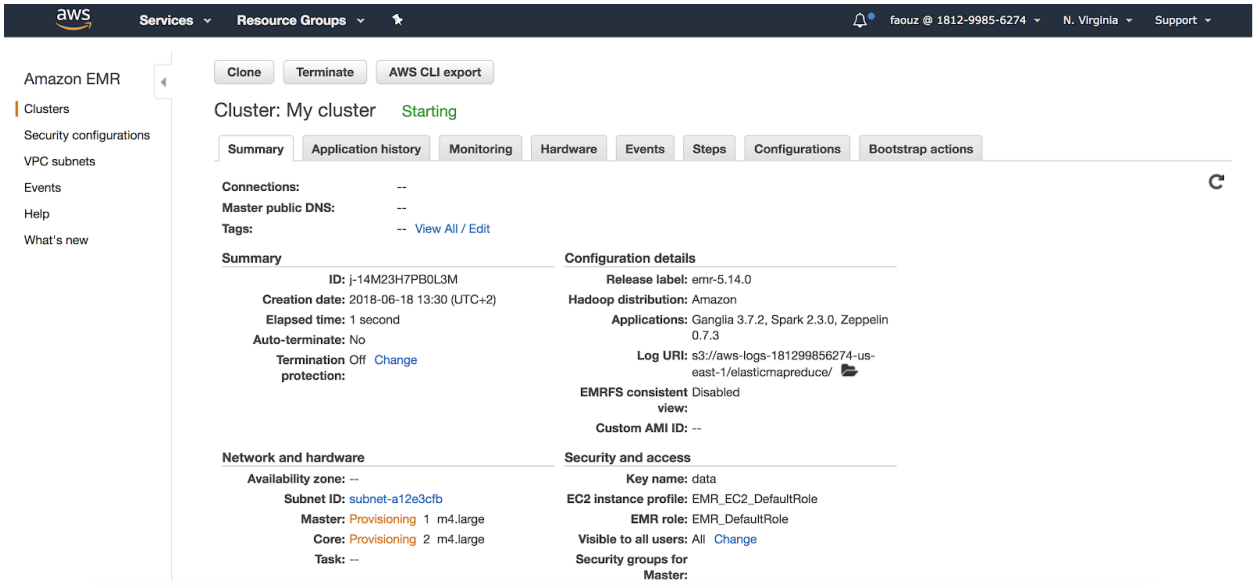
- Sélectionnez Spark dans la partie “Applications”



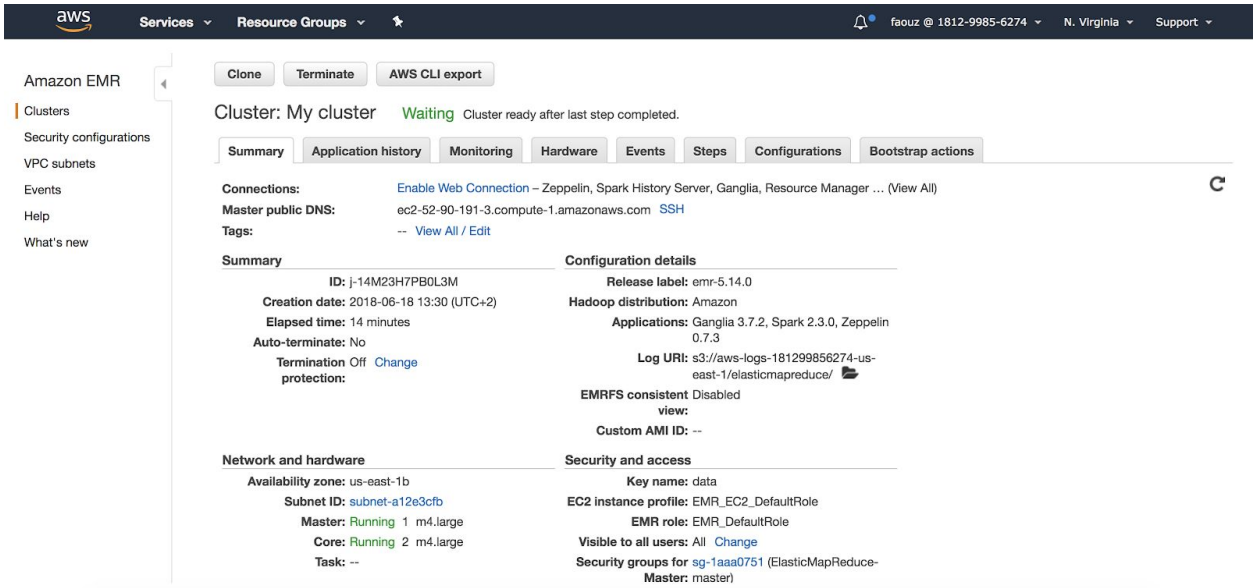
- Puis, dans la partie type d'instance sélectionnez **m4.large**, au nombre de 3.



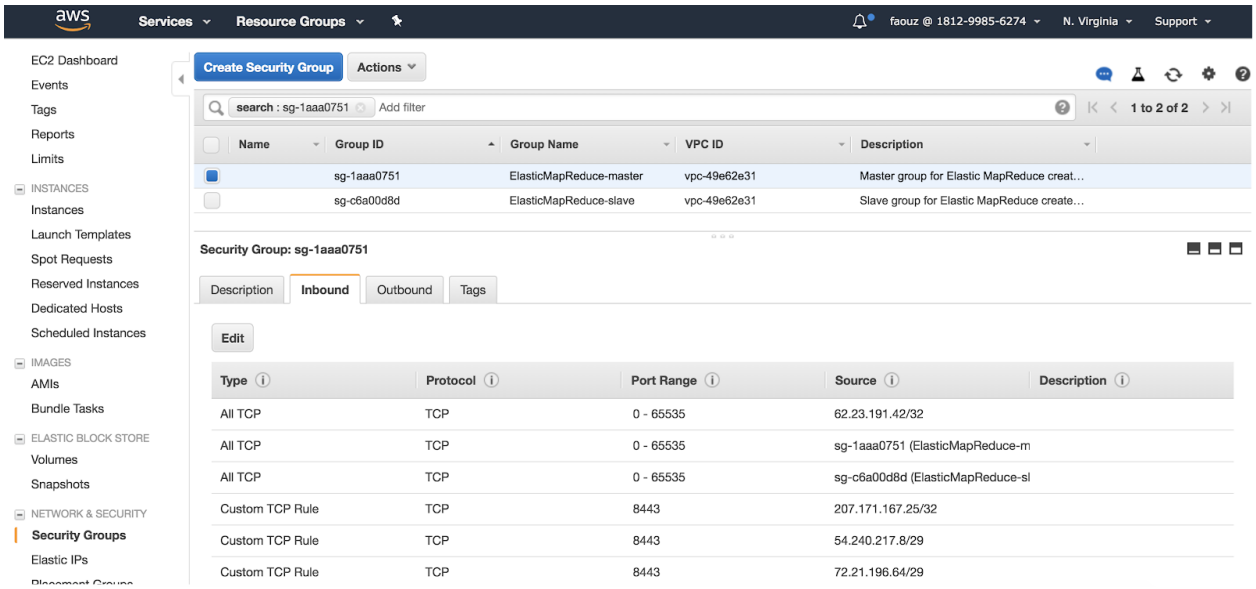
- Attendez quelques minutes pour que votre cluster démarre.



- Une fois démarrer, le cluster se met en attente

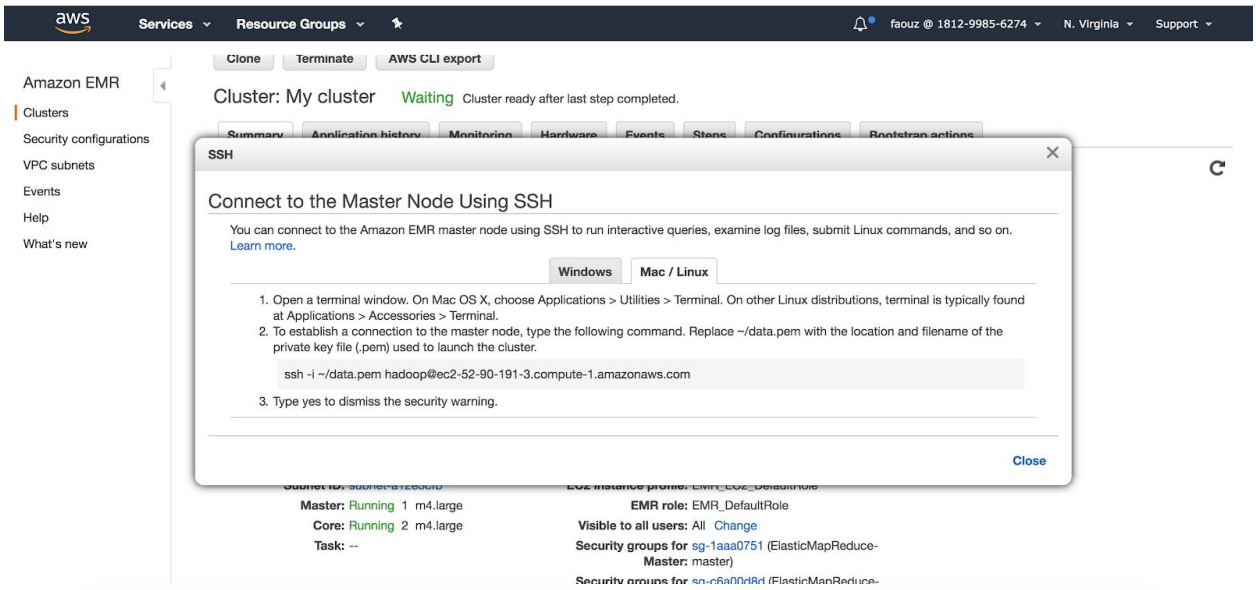


- Cliquez sur **Security groups for Master,** puis dans l'onglet Inbound



Cliquez sur edit, puis ajouter une règle pour autoriser tous les ports (0 - 65535) à votre adresse IP.

- Enfin, cliquez sur le lien SSH pour avoir les instructions pour pouvoir s'y connecter en SSH.



En copiant cette commande dans votre terminal vous pourrez vous-y connecter pour utiliser le client **pyspark** ou **spark-shell**.