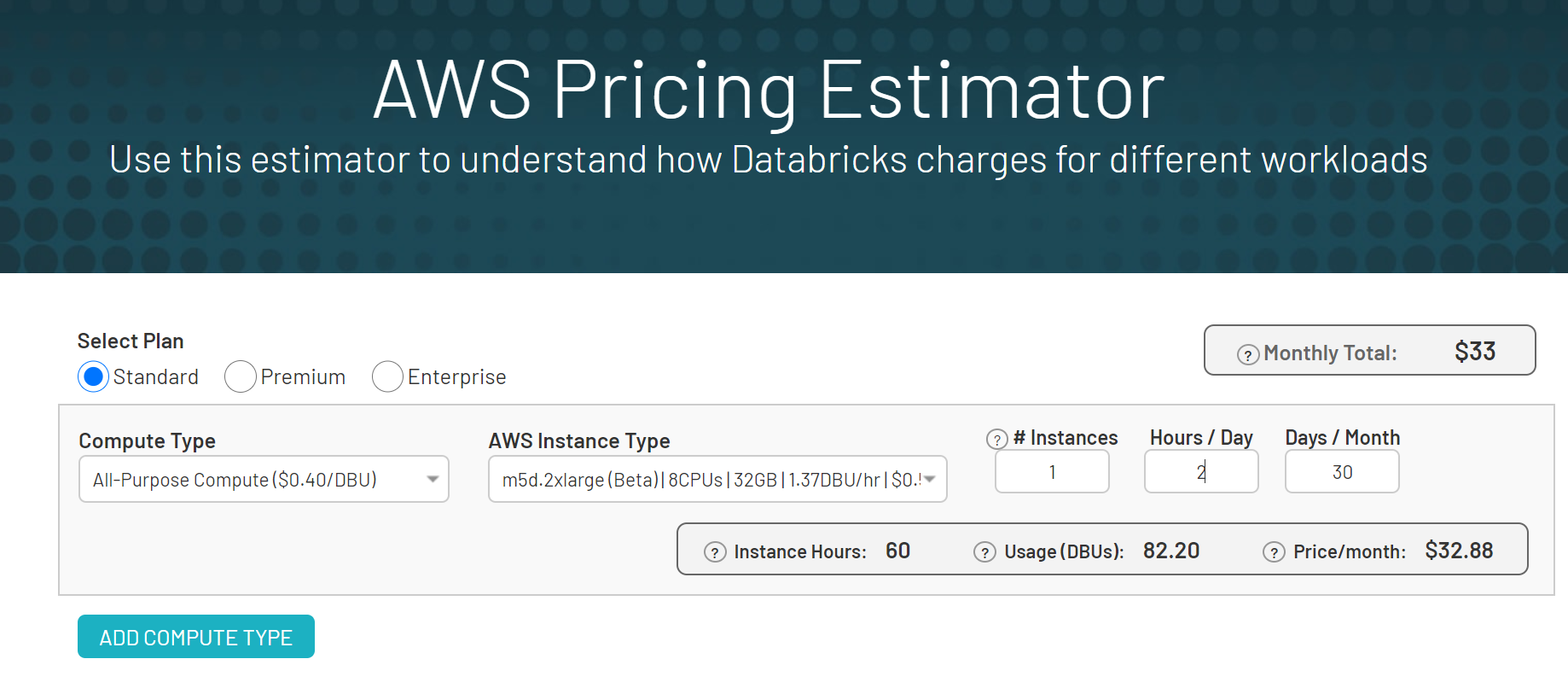
**Reasons for running the scripts on local machine**

We decided to have the scripts run on a local machine (automatically using task scheduler without any manual dependency) because storing the scripts and running them on the external servers would be expensive.

The 2 other options we had were:

1. Store the scripts on an instance of the AWS EC2. But the problem with this was the fact that the EC2 instance needed to be logged in using a local machine and would need manual dependency to that. Since one of our goals for this project was to make it E2E automated, this option was ruled out.

2. Another option that we had was to use Azure data bricks. In this scenario we could spin up a VM on Azure and install python to run scripts and automate the scripts. But the reason we did not go ahead with this option was because of the cost involved. Since we are already paying around 20 euros a month for our Azure subscription, we did not want to spend more on it.



Considering we only keep the VM running for 2 hours a day until all the pipelines run, it would cost us about 33 euros a month for the subscription. This is just the VM cost, we need to pay extra for the storage costs. Also, if the VM needs to be spun up and shut down every day, that will be a manual process that cannot be automated.

Keeping all these in mind, we decided to create cron jobs on our local machine. and run task scheduler jobs for our python scripts. As a result, we do not have any manual dependency and this also makes it easier to move it to an external server in the future if the project is scaled.