Executing Data Ingestion Framework from AWS

Approach:

Executing DI Framework from AWS environment by placing framework code in EC2 compute node and run from there itself.

AWS Components:

For this approach following AWS components required

* AWS account
* Ec2 cluster

Estimated Cost:

* Ec2 with 1 instance type c5.xlarge cost is $66 per month.

Pros:

* Install python and spark.
* Easy approach with controlled flow

Cons:

* Cluster always up and run
* Manual execution than EC2 Steps automation

Automation need another support.

**Other Storage areas cost: -**

**DynamoDB: -**

Estimated Cost:

* DynamoDB with 1Instance 100 write capacity units cost is $1893per month.

For 1TB storage with 100DTU’s/hr, 1 instance reserved capacity and 1 instance for OnDemand.

**Redshift DB: -**

Estimated Cost:

* Redshift DB with ds1.xlarge-2TB HDD cost is $1846 per month

For 2TB storage with 1000DTU’s/hr, 1 instance reserved capacity and 1 instance for OnDemand.